Arizona State University
On Behalf Of
PACE: Producing Algae for Coproducts and Energy Consortium

Evaluation of Genetically-Modified *Chlorella sorokiniana* in Open Ponds for Production of Biofuel Feedstock and High Value Co-products

---

TSCA Environmental Release Application (TERA)

ID # PACE03

Date of submission
September 2018

Table of Contents

List of Figures and Sequences

List of Figures and Sequences (cont.)

List of Tables

List of Attachments

*All scientific articles have been combined into a single PDF, "ScientificArticles.pdf"

*All SOPs have been combined into a single PDF, "SOPs.pdf"

Certification Statement -

In this document we describe the construction of 1 strain, PACE_Cs1412_SNRK2. This strain is based on the recipient microorganism, *C. sorokiniana (DOE1412)*.

Additionally, I certify that to the best of my knowledge Arizona State University in collaboration with the Colorado School of Mines as part of the PACE Consortium project DE-EE0007089 intends to perform only the activities outlined in this submission. The information contained within this submission is truthful and accurate as of the date of this submission.

Sincerely,

John A. McGowen

Dr. John A. McGowen

Director of Operations and Program Management

Arizona Center for Algae Technology and Innovation

Date: 8/16/2017

A. RECIPIENT ORGANISM CHARACTERIZATION

C. sorokiniana is a unicellular, green alga that has been used as a model organism for photosynthesis studies and in various practical applications in agriculture, biotechnology, and as a food additive. Members of the Genus *Chlorella* have small, spherical or ellipsoidal cells and are globally distributed; naturally occurring on soil and in freshwater. Strains of *C. sorokiniana* are generally observed to be non-flagellate cells but contain a vestigial flagellar apparatus. Sexual reproduction has not been reported or observed to date, thus, reproduction is achieved by producing non-motile asexual autospores. Most *Chlorella* have a polysaccharide cell wall containing a sporopollenin-like substance that occurs in the walls of the pollen grains of higher plants.

Strains of *C. sorokiniana* belong to the Order *Chlorellales* in the Class *Trebouxiophyceae*. As demonstrated in the literature, *C. sorokiniana* can be distinguished from other Trebouxiophyceae using the ITS2 gene sequence (Neofotis, 2016), and by comparison of the chloroplast genomic DNA (Lemeiux, 2014). Different *Chlorella* species and different *C. sorokiniana* strains of the same species can also be readily resolved by comparing the nuclear genome sequence (Barry, 2015) or using the rDNA region including the 3' end of the 18S gene, the ITS1 region, the 5.8S gene, the ITS2 region, and 5' end of the 28S gene (Figure 1; Bock, 2011; Neofotis, 2016).

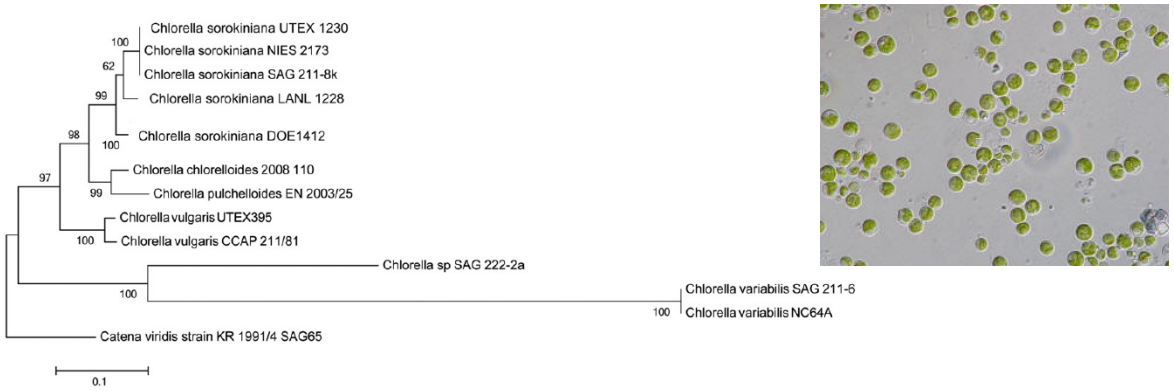


Figure 1: *Chlorella* phylogenetic tree based on the rDNA region including the 3' end of the 18S gene, the ITS1 region, the 5.8S gene, the ITS2 region, and 5' end of the 28S gene. Adapted from Neofotis et al, 2016, and a micrograph (inset) of *C. sorokiniana.*

The recipient strain for this project will be *C. sorokiniana* DOE1412. This organism can be identified by running a whole cell approach to PCR with the specific primers developed for allowing discrimination from other *Chlorella* sp., even specific strains within species. The *C. sorokiniana* 1412 specific primers are a) FWD 5' GCGAAGAAGAAAATGTAAACTTATTAG 3' and b) Rev 5' CCATTCCAGTAATTGCTAAATCA 3'

B. SUBJECT ORGANISM CHARACTERIZATION

B.1 Taxonomies of the Subject and Donor Microorganisms

B.1.a. If genetic manipulation is so extensive that a subject microorganism might be more appropriately assigned to a different taxon than the recipient parent, the submitter should provide equivalent support for the designation that applies to the final construct as that provided for the designation of the recipient parent. Such support should apply to the microorganism intended to be used as the final production strain.

The taxonomy of the subject microorganisms is identical to the recipient microorganism. The only

difference is that in the subject microorganisms two foreign genes are expressed under a relatively strong native promoter. SNRK (SNF related kinase) play a key role in sugar metabolism in pant and animal kingdom and control multiple growth and metabolic processes. We have utilized SNRK 2 gene to improve photosynthetic efficiency and biomass in the recipient organism. The regulatory elements used to express SNRK gene is the psaD (a photosynthesis-related gene) and actin promoters and terminators, both of which are endogenous to the recipient microorganism. The purpose of this field experiment is to: 1) evaluate the translatability of GM phenotypes from a lab to an outdoor setting, and 2) to compare the resistance of the GM strains to biotic (bacterial) and abiotic (diurnal temperature and solar insolation) in the subject microorganisms versus the recipient (wild type) strain. The genetic modifications made to the subject microorganism are described in the sections that follow.

B.1.b. A generic name for the subject microorganism is needed if the name of the subject microorganism is to be considered as Confidential Business Information (CBI). Any trade name under which the subject microorganism might be marketed should also be supplied.

From this point forward in the document, the organisms will be referred to by strain names as denoted in the table below:

Table 1: Strains included in this TERA application.

Strain Name	Vector Introduced	Genes expressed	Notes
C. sorokiniana DOE1412	N/A	N/A	Recipient Microorganism
PACE_Cs1412_SNRK2	SNRK2_PACE_ plasmid	SNRK 2 from *Picochlorum soloecismus*	Subject Microorganism

B.1.c. Taxonomic characterization of donor microorganisms, which contribute intergeneric DNA to the subject microorganism, or provide intrageneric DNA that may affect the expression or stability/transfer of the intergeneric DNA is needed. Characterization could include identification of a genus, species, and strain designation for each donor microorganism. Generally, support information for these taxonomic designations need not be provided in detail as suggested for the recipient microorganism under item IV.A. Taxonomic information on the microorganisms from which the nucleic acid sequences were first isolated may also be helpful, if their taxonomies differ from that of the donor microorganisms used to construct the subject microorganism.

The intergeneric gene used to develop the strains in this TERA, SNRK 2 gene from *Picochlorum soloecismus* strain, which is a genus of green algae in the class Trebouxiophyceae. The gene was synthesized in its native state (only the coding regions) without codon optimization and cloned into the PACE *Chlorella* plasmid vector shown in Figure 2 by Genewiz (https://www.genewiz.com/en).

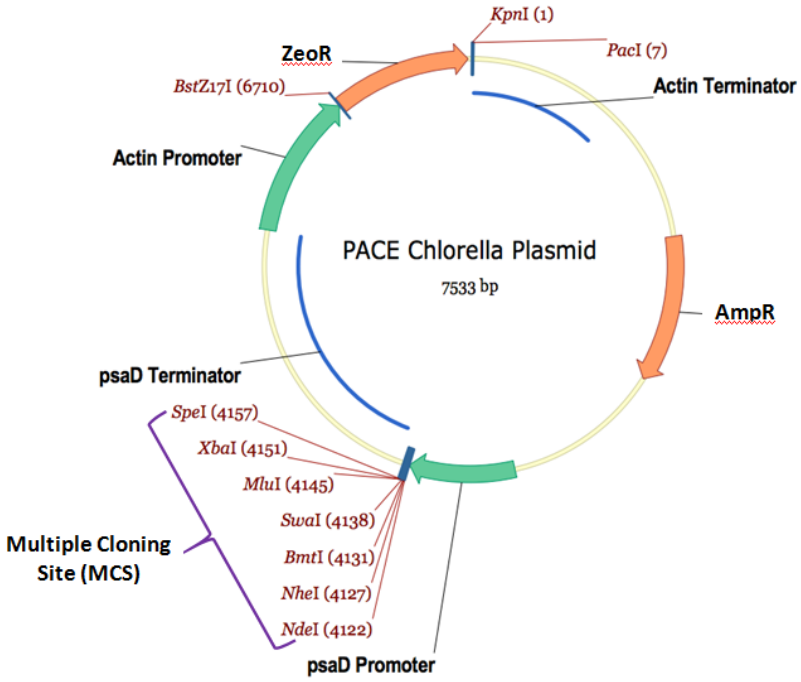


Figure 2: Map of PACE *Chlorella* plasmid developed as a vector for introducing target genes into the recipient organism, *C. sorokiniana UTEX* 1412, showing the location of the gene cloning MCS and the representative restriction enzymes located therein. AmpR = Ampicillin resistance gene; ZeoR = Zeocin resistance gene; psaD = native *C. sorokiniana* photosynthesis-related gene promoter/terminator pair.

The PACE vector was developed by researchers at the New Mexico Consortium to introduce genes of interest into *Chlorella sp.* and is built on the *E. coli* vector backbone from plasmid pSL18. It contains:

1) An *E. coli* plasmid origin of replication and an ampicillin resistance (AmpR) gene for propagation in *E. coli*.

2) The *Streptoalloteichus hindustanus* Sh *ble* gene, which confers zeocin (ZeoR) resistance, under the control of the *C. sorokiniana* actin promoter/terminator The sh-lbe gene, which confers resistance to both zeocin and bleomycin, was first isolated from *Streptoalloteichus hindustanus* (https://doi.org/10.1016%2F0014-5793%2888%2980665-3). The gene was codon optimized for Cs1412 expression, however, the protein sequence remained unchanged.

3) An empty multiple cloning site (MCS) for insertion of the gene of interest; and the psaD promoter/terminator pair flanking the MCS. This promoter was chosen due to its relatively high expression as a native photosynthesis-related gene promoter in *C. sorokiniana*.

The vectors described in section B.2 were introduced into the recipient microorganism by electroporation transformation technique using the transformation protocol in Section B3.m.

B.2 Final Constructs

Provide an illustration of the final construct, which is in the subject microorganism identified in item IV.B.1, and again note the taxonomy of the subject microorganism. Provide a detailed legend to support the final construct illustration.

The final vector construct used to create the subject microorganism is shown in Figure 3. The full sequence is shown in Figure 3A. The genetic elements of this construct are described in Table 2

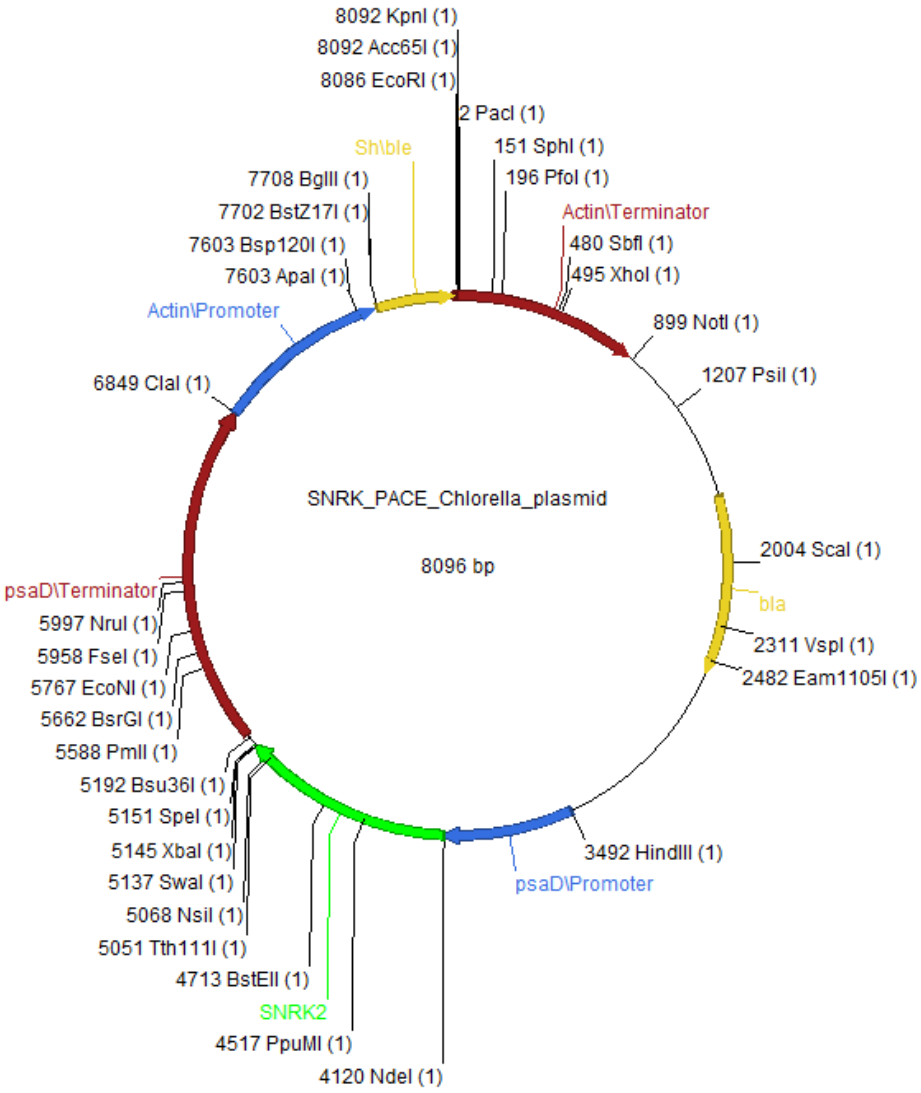


Figure 3: Map of SNRK_PACE_Chlorella_Plasmid showing the final transformation plasmid construct

Figure 3A: Sequence of PACE_ SNRK_PACE_Chlorella_plasmid (map shown in Figure 3)

cttaattaagagctcctcgggccggcagtagttgctacctgggagtggtcggcgcggtgtgcacttgaagcaggggcggcttgcgtagtgcag
cagcgtgggaggaaattgcaaagactgtgcagcccacctcatctagcgcctgggatcgcatgcacacttcacatgctggcaatggtctaggc
cctgccctttcctggaggctgctctgccatacaccttgcctgacattacagcatggactgttgcatcgggtcgacagcgcacatgcaagccgag
ccacaccgctgcattcgagcatcaggccagcaccatgggctgcccaggcaggcagcctagcgtccgcccagcttctgccgcagggcttgca
gggcatccggcccgctggcctcccggcccagcgcagtggcaaccacagtcatgccacggccgcccgctttcttgcgctgcttggctgtcagg
gccttgttgagtgccgcctgcagggccgccgctcgagggtcaacaaccttctttttggcggccttcttcttctgctgcttggcgcccgatgagcccg
ctgtgccagctgcgcccgctgcactcgcccggctgcggctgtccgactgccgctgagccagagccacagcctggctcacagtctgcactgtgc
gaggctccggcacaggtgccgctgcctcgcctgccgctgccgcaccctctgctgctgccgccgctgctttcttgcgcttcttgccaccggctgctg
cgggggccgaagggtcgactgccgccatgactgctgctgctgcgccgccggtgaaccacttgcggctgaacttgggctgcagcggctcggc
cagcagcgctgccagctgtgcacgcacatctgctgccgtctctgcagcacagcattgtgcggccgccaccgcggtggagctccaattcgccct
atagtgagtcgtattacaattcactggccgtcgttttacaacgtcgtgactgggaaaaccctggcgttacccaacttaatcgccttgcagcacatc
ccccttcgccagctggcgtaatagcgaagaggcccgcaccgatcgcccttcccaacagttgcgcagcctgaatggcgaatggaaattgtaa
gcgttaatattttgttaaaattcgcgttaaatttttgttaaatcagctcatttttaaccaataggccgaaatcggcaaaatcccttataaatcaaaaga

atagaccgagatagggttgagtgttgttccagtttggaacaagagtccactattaaagaacgtggactccaacgtcaaagggcgaaaaaccg
tctatcagggcgatggcccactacgtgaaccatcaccctaatcaagtttttttggggtcgaggtgccgtaaagcactaaatcggaaccctaaagg
gagcccccgatttagagcttgacggggaaagccggcgaacgtggcgagaaaggaagggaagaaagcgaaaggagcgggcgctaggg
cgctggcaagtgtagcggtcacgctgcgcgtaaccaccacacccgccgcgcttaatgcgccgctacagggcgcgtcaggtggcacttttcgg
ggaaatgtgcgcggaacccctatttgtttatttttctaaatacattcaaatatgtatccgctcatgagacaataaccctgataaatgcttcaataatatt
gaaaaaggaagagtatgagtattcaacatttccgtgtcgcccttattccctttttgcggcattttgccttcctgtttttgctcacccagaaacgctggtg
aaagtaaaagatgctgaagatcagttgggtgcacgagtgggttacatcgaactggatctcaacagcggtaagatccttgagagttttcgccccg
aagaacgttttccaatgatgagcacttttaaagttctgctatgtggcgcggtattatcccgtattgacgccgggcaagagcaactcggtcgccgca
tacactattctcagaatgacttggttgagtactcaccagtcacagaaaagcatcttacggatggcatgacagtaagagaattatgcagtgctgcc
ataaccatgagtgataacactgcggccaacttacttctgacaacgatcggaggaccgaaggagctaaccgcttttttgcacaacatgggggat
catgtaactcgccttgatcgttgggaaccggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctgtagcaatggcaa
caacgttgcgcaaactattaactggcgaactacttactctagcttcccggcaacaattaatagactggatggaggcggataaagttgcaggac
cacttctgcgctcggcccttccggctggctggtttattgctgataaatctggagccggtgagcgtgggtctcgcggtatcattgcagcactggggc
cagatggtaagccctcccgtatcgtagttatctacacgacggggagtcaggcaactatggatgaacgaaatagacagatcgctgagataggt
gcctcactgattaagcattggtaactgtcagaccaagtttactcatatatactttagattgatttaaaacttcatttttaatttaaaaggatctaggtgaa
gatccttttttgataatctcatgaccaaaatcccttaacgtgagttttcgttccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgag
atccttttttctgcgcgtaatctgctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgccggatcaagagctaccaactctttttc
cgaaggtaactggcttcagcagagcgcagataccaaatactgttcttctagtgtagccgtagttaggccaccacttcaagaactctgtagcacc
gcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggcgataagtcgtgtcttaccgggttggactcaagacgatagttacc
ggataaggcgcagcggtcgggctgaacgggggggttcgtgcacacagcccagcttggagcgaacgacctacaccgaactgagatacctac
agcgtgagctatgagaaagcgccacgcttcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcggaacaggagagc
gcacgagggagcttccagggggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgatgctcgtca
ggggggcggagcctatggaaaaacgccagcaacgcggccttttacggttcctggccttttgctggccttttgctcacatgttctttcctgcgttatcc
cctgattctgtggataaccgtattaccgcctttgagtgagctgataccgctcgccgcagccgaacgaccgagcgcagcgagtcagtgagcga
ggaagcttttcagcaggttattggttcatgcacggtcatgttcctggtcgcttgtatctgcgatctgcaacgatatccatgatgacacgcaagtccac
gcattgagcagctccgtggtgccacagctggtccagccttgctacaagtgcgggtagagcaacagcacagccgtggcggctggaatttgcca
tggcctgtccatatctaccaaccagcgcgggatctgtgaagggcatgtgagcggcgccgcacgggcggcggctgccgccaaccggctgtcg
ccgcctgctgcagcgggtcgcctgacccccttgcagtgacttgctgttcacatcaagagcaatgaatgaatgaatgccctgccgcggtgaatgcccg
cagctgctgccccaggcagtcgggggacggcaccccagccggcaggtcgcccggcgcggattgccggcggctttctgaatgcaggcact
gccgataaagtggttgggctcccacacagcactttgcagcgccaccctgctgagcatttcaccgggtcacagcaggatcgtcagcgagaaag
gctccccccccagcgaccgctcggccgctgctcccacccccatatagccccagccgaccaggacagaggccaaacCATATGGCTAG
CGCTAGCATGGATATGCTACCTGAACGTTACGAGCTGATAAGAGACATTGGCAGTGGAAATTTTG
GTGTAGCCAAGTTAATGCGTGACCGAAAAACCGGCGAGCAGGTTGCTGTGAAGTTCATCGAGAG
AGGAGAAAAGATCGATAAGAATGTCGAGCGTGAAATTGTCAATCACCGTCAACTTTCTGGACACC
CCAACATAATACGGTTCAACGAGGTCTTTCTCACCTCAACTCATTTAGGCATATCTATGGAATATG
CTTCAGGAGGAGAATTATTCGATCGTATCGTAAAGGCTGGGAGGTTTTCAGAGGATGAGGCTAG
ATATTTCTTTCAGCAGTTGATATCAGGAGTAAACTGGTGCCACAAAGAGGGTGTATGCCATAGGG
ACCTAAAATTGGAAAATACGCTACTAGATGGTAGAGCTGCACCAAGATTGAAAATATGCGATTTT
GGATATTCTAAATCGGCTATTTTTGACTCGCAGCCGAAATCTACCGTAGGTACGCCCGCCTATAT
AGCGCCTGAAGTTCTATCCAGAAAGCAATATGCGGGGGAGATTGCTGATGTTTGGTCATGTGGG
GTAACCTTGTACGTTATGCTAGTGGGTGCATATCCCTTTGAAGATGCAAACGATCCTAGAAACTTT
AGAAAAACCATCCAACGGATCATGGGTGTGAAATACTCGTTCCCGAGCAATCTTCCGCTTAGTAA
AGAGTGTCATGATCTGATGAGTAGGATTTTTGTTGCTAATCCATCCAAAAGAATTTCATTGACGGA
GATTAAGGCACATCCATGGTTTTTGAAGAACTTGCCTCCAGAATTAGATGAGAAAGAAATTATCG
GAGAAGTGAGATCAAAAAACTCTGATATACAGCCAGTAGAGGAGATCAAGAGACTTGTTGCAGAT

GCAAGGATGAGACCAAGTCAAGCTGGAATGCATACCTTCAACGAAGATGACTACATGGACGGCG
ATATAGATATGGAGACAGAAATTATGGATTAAATTTAAATTCTAGAACTAGTgctggccactcggcccagca
gcttggcagcgctggcctgaggggcgcctggccatggcgtcctgagcgctggcgcagcttgcagccaccagggccgccggtgcggcgtgg
ccggcggcctgccggccgcagctgcagctgcggcgcgcggccggcagcggcggcgaaggggtggtgggcggcgctgcggcggcgtgg
ggacggctcccagattagtgccgtgccctcgtgctgcgcccgcccagcactcaaccgccgccgccctcacccacccccctgagaagaactaa
ccaaccctcacagtttctagtgccacagcgtgtacgatacgtcctggcaattatgtctgtctcattaaatgctgctgctgtttttgttgcaaaccgcta
gatttgcacacccctctctcccggccgtgtaaggctgaacccaacacgtggcaccggggattgttgcgagcctggctggggaagacatgggc
agctggttgagatgttgaattacagctgtacatcagtttgcgtgcattgtgtccgaacctggagcactgcaaggcactggcggctggaagtggct
ggtcagttcactcttcattcaccagcaccacctcctcctgcaccagggcagcctcccctccgcagctgccgccgcagctgaacacagccacg
gtgagccactcccagctgtcgggcggctgcagggaggctgccgctgctgctccgccctcctcagctgccagccattcggctgcctcggtcaaa
gcggggatgaggggtgtcatcagctgcagctcaaacacgcgcggccggccgcaggcggcacagggtgcgggagcgggctgctcgcgag
tcggccacagcggtgcaccgccaaagctgcagggtgaaggggccagaggacagcagagaagtcggttttgggtcacttcagtagtcacca
gcgggcagcaaaactgacctccagccctcttgcttgcactgaccacccctttggcggccccgctcgcctgtagcgcgcgcactggtcggggca
gcggcccagctgctttgcaaacttgaagaaggcggctcccacgcccgcccgcgtgccgccggtgggcgccagcacagcgtcctcttcatag
ccttctcctgaccaggactcagggccgccctccacgctgggcgcacatgctgtgcccgctgtggtcgccatggcggcgcccgcctcgccagct
gctgccgcctctgcctcgtatcgagccatcagctgctgcaggtgctcctgttcggagcgctgctgcctgccgctgctggcgctgcactgcgactc
cggctcggcatacaggtagaaggcggggaggctgggctgggttgggtcatacgctggcgctgccccagctgctgccacctcgctgctgcgca
gctccccctgctgctgcggctgctgcttgccggctggcttggcagctggggcagcactgcccaccgcttccagcgctgcgttgaggtctccaaa
gtcaaacggagctgccgcactgtccgtcacatcagcgcccatcccccagtcatcgctgcctcctgcaccccagtcatctgcaccgcccatacc
ccagtcatcggctgcaggcggcttctctgccggcggcgctgcctgctgctgcagctgctgcgctgtcggcggctgctcgctgctggctggagtgc
tgggcagggtatcgatggcggcagccagctcaagtgtgtgtggctctgggctgctgccggtgttgaccgccgctgcgccgccgcccggcagc
cagtgcaccagcagctgcgggccggcggacaggcacttgacatacagcggcgaacggctgcctgcaagcgggaatcggttgaggtcagc
tacggtgcctcgtgacggcagccagcagcgatcagcaagccacggctgggagccatggcagccacccgaccaacacccaacaacccag
cggcaacccaggtcagcagccgctgcacgcgccttgcgagtcttggtagcgaaaggcgtaggaaccctccagagattcccacccagaggtt
cccacttcctcacggtcggcgccccagtctgcacatgtggatgggggtggggcaattagtagcgaggatcggttaggcagggctggcacgc
gagggcaaagggccgaacagggctcttggaaagctgatcatgcacccagcaagtgcaatcggccactgtgattggctgcttctctaaacatct
gatcctgatccagcagacaaagccaagatccggcctctcaccggcaccctcgtcctcggcggcagggccggcggccaccagcttgtagcct
tcggccagcaggaaggcgtgcaccgcaaacgccagcttgtcggcgtcggacctggcgagcatgagaagatacaacttggttgacgatgtgc
aggtgtgctggactggccgaggggctggacggccgaggggcccgcaggcggctggcggcgatggcggcggcgcggcggcggtgccag
cagcagcgctttgcggcccggcaccttgctgacctgaagttggggcgcggtatacagatctatggccaagctgacctccgccgtgcccgtgct
gaccgcccgcgacgtggccggcgccgtggagttctggaccgaccgcctgggcttctcccgcgacttcgtggaggacgacttcgccggcgtgg
tgcgcgacgacgtgaccctgttcatctccgccgtgcaggaccaggtggtgcccgacaacaccctggcctgggtgtgggtgcgcggcctggac
gagctgtacgccgagtggtccgaggtggtgtccaccaacttccgcgacgcctccggccccgccatgaccgagatcggcgagcagccctggg
gccgcgagttcgccctgcgcgaccccgccggcaactgcgtgcacttcgtggccgaggagcaggacgaattcggtac

Table 2: Genetic elements contained in SNRK_PACE_Chlorella_Plasmid.

Vector	Element Name	Element Type	Length (bp =base pairs)	Location on vector (bp #s)	Description
SNRK_PACE_Chlorella_Plasmid	Multiple cloning site	Linker site	15	1-15	Multiple cloning site PacI, ACC65, Kpn
	Actin Terminator	Regulatory	883	16-898	Native *C. sorokiniana* actin gene terminator
	Backbone sequence	Nucleotides	32	899-930	*E. coli* plasmid pSL18 vector backbone
	T7 Promoter	Regulatory	19	931-949	*E. coli* plasmid pSL18 vector backbone
	Backbone sequence	Nucleotides	6	950-955	*E. coli* plasmid pSL18 vector backbone
	M13 fwd	Regulatory	17	956-972	*E. coli* plasmid pSL18 vector backbone
	Backbone sequence	Nucleotides	140	973-1112	*E. coli* plasmid pSL18 vector backbone
	F1 Ori	Replicating	456	1113-1568	*E. coli* plasmid pSL18 vector backbone
	Backbone sequence	Nucleotides	26	1569-1594	Ampicillin resistance gene promoter on *E. coli* plasmid pSL18 vector backbone
	AmpR gene promoter	Regulatory	105	1595-1699	Ampicillin resistance gene promoter on *E. coli* plasmid pSL18 vector backbone
	AmpR gene	Selection Marker	861	1700-2560	Ampicillin resistance gene (codes for beta-lectamase) on *E. coli* plasmid pSL18 vector backbone
	Backbone sequence	Nucleotides	170	2561-2730	*E. coli* plasmid pSL18 vector backbone
	Ori	Replicating	589	2731-3319	*E. coli* origin of replication on pSL18 vector backbone
	Backbone sequence	Nucleotides	178	3320-3497	*E. coli* plasmid pSL18 vector backbone
	psaD Promoter	Regulatory	622	3498-4119	Native *C. sorokiniana* promoter for the psaD gene coding for subunit ii of photosystem I reaction center
	SNRK	Intergenic Gene	1037	4120-5156	Synthetic gene sequence of the sucrose non-fermenting protein kinase gene from *Picochlorum solecismus*
	Multiple cloning site	Linker site	45	5157-5201	Multiple cloning site SpeI,Bsu36I
	psaD Terminator	Regulatory	1647	5202-6848	Native *C. sorokiniana* terminator for the psaD gene coding for subunit ii of photosystem I reaction center

	Multiple cloning site	Linker site	6	6849-6854	Multiple cloning site ClaI
	Actin Promoter	Regulatory	847	6855-7701	Native *C. sorokiniana* actin gene promoter
	Multiple cloning site	Linker site	12	7702-7713	Multiple cloning site BglII, BstZ17
	ZeoR	Selection Marker	372	7714-8085	Sh ble gene from *Streptoalloteichus hindustanus*, which confers zeomycin resistance
	Multiple cloning site	Linker site	11	8086-8096	Multiple cloning site EcoRI,Acc65I,KpnI

B.3 Construction of the Subject Microorganism

B.3.a. A brief summary of the construction strategy should be presented. The summary should indicate why the genetic manipulations were done and their effect(s) relative to the recipient microorganism. This summary should allow the Agency to follow the manipulations done, beginning with the original sources of the introduced DNA, to arrive at the final construct

The plasmid pSL18, which can be replicated in competent bacterial cells and which can drive expression of genes of interest when randomly integrated into the *Chlorella* genome, was used as a backbone for the synthesis of the final PACE vector containing SNRK2 gene coding sequences (CDS) cloned into the MCS. The Non codon optimized gene sequence (CDs) was synthesized and cloned between psaD promoter and terminator by Genewiz (www.genewiz.com) using the NdeI and SpeI unique restriction sites in the MCS (Figure 2). Similarly, Genewiz was responsible for the synthesis and cloning of the ZeoR resistance gene (replacing the original paramomycin resistance gene) between the actin promoter and terminator using the BglII and EcoRI unique restriction sites.

The final construct shown in Figure 3 was built on the backbone of the *Clamydomonas reinhardtii* pSL18 plasmid shown in Figure 4 with the sequence immediately following in Figure 4A. The first genetic modification to pSL18 was to replace the *C. reinhardtii* psaD promoter/terminator pair with the *Auxenochlorella protothecoides* (formerly known as *Chlorella protothecoides* (CP)) promoter/terminator pair. These *Chlorella* specific genes for the CP Psad promoter/terminator pair added to pSL18, were synthesized and cloned into the vector by Genewiz. Figure 5 shows the annotated sequence of the *Chlorella* specific promoter/terminator pair insert. The map of the resultant plasmid, pSL18CP (also known as pCp) is shown in Figure 6 with the sequence immediately following in Figure 6A. The pSL18CP vector was sequence confirmed by Genewiz using the primers listed in Figure 7.

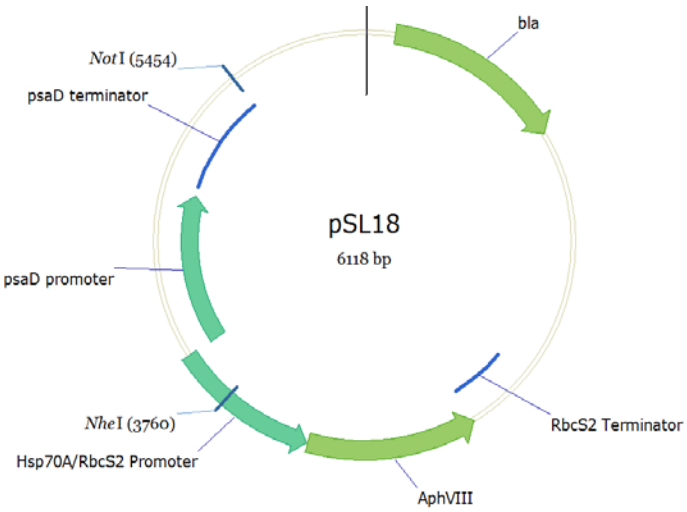


Figure 4: Map of pSL18 plasmid which forms the backbone of the *Chlorella* plasmid developed as a vector for introducing target genes into the recipient organism, *C. sorokiniana* DOE1412 showing the NheI and NotI restriction sites used to insert a *Chlorella* promoter/terminator pair. bla = Ampicillin resistance gene; AphVIII = Paromomycin resistance gene; psaD = native *Clamydomonas reinhardtii* photosynthesis-related gene promoter/terminator pair.

Figure 4A: Sequence of pSL18 plasmid (map shown in Figure 4)

```
cgtcaggtggcacttttcggggaaatgtgcgcggaacccctatttgtttatttttctaaatacattcaa
atatgtatccgctcatgagacaataaccctgataaatgcttcaataatattgaaaaaggaagagtatga
gtattcaacatttccgtgtcgcccttattccctttttgcggcattttgccttcctgtttttgctcacc
cagaaacgctggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtgggttacatcgaactgg
atctcaacagcggtaagatccttgagagttttcgccccgaagaacgttttccaatgatgagcactttta
aagttctgctatgtggcgcggtattatcccgtattgacgccgggcaagagcaactcggtcgccgcatac
actattctcagaatgacttggttgagtactcaccagtcacagaaaagcatcttacggatggcatgacag
taagagaattatgcagtgctgccataaccatgagtgataacactgcggccaacttacttctgacaacga
tcggaggaccgaaggagctaaccgctttttgcacaacatgggggatcatgtaactcgccttgatcgtt
gggaaccggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctgtagcaatggcaa
caacgttgcgcaaactattaactggcgaactacttactctagcttcccggcaacaattaatagactgga
tggaggcggataaagttgcaggaccacttctgcgctcggcccttccggctggctggtttattgctgata
aatctggagccggtgagcgtgggtctcgcggtatcattgcagcactggggccagatggtaagccctccc
gtatcgtagttatctacacgacggggagtcaggcaactatggatgaacgaaatagacagatcgctgaga
taggtgcctcactgattaagcattggtaactgtcagaccaagtttactcatatatactttagattgatt
taaaacttcatttttaatttaaaaggatctaggtgaagatccttttgataatctcatgaccaaaatcc
cttaacgtgagttttcgttccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatc
ctttttttctgcgcgtaatctgctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgc
cggatcaagagctaccaactctttttccgaaggtaactggcttcagcagagcgcagataccaaatactg
ttcttctagtgtagccgtagttaggccaccacttcaagaactctgtagcaccgcctacatacctcgctc
tgctaatcctgttaccagtggctgctgccagtggcgataagtcgtgtcttaccgggttggactcaagac
gatagttaccggataaggcgcagcggtcgggctgaacgggggttcgtgcacacagcccagcttggagc
gaacgacctacaccgaactgagatacctacagcgtgagctatgagaaagcgccacgcttcccgaaggga
gaaaggcggacaggtatccggtaagcggcagggtcggaacaggagagcgcacgagggagcttccagggg
gaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgacttgagcgtcgattttgtgat
gctcgtcaggggggcggagcctatggaaaaacgccagcaacgcggccttttacggttcctggccttttt
gctggccttttgctcacatgttctttcctgcgttatcccctgattctgtggataaccgtattaccgcct
ttgagtgagctgataccgctcgccgcagccgaacgaccgagcgcagcgagtcagtgagcgaggaagcgg
```

aagagcgcccaatacgcaaaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgaca
ggtttcccgactggaaagcgggcagtgagcgcaacgcaattaatgtgagttagctcactcattaggcac
cccaggctttacactttatgcttccggctcgtatgttgtgtggaattgtgagcggataacaatttcaca
caggaaacagctatgaccatgattacgccaagctcgaaattaaccctcactaaagggaacaaaagctgg
tacccgcttcaaatacgcccagcccgcccatggagaaagaggccaaaatcaacggaggatcgttacaac
caacaaaattgcaaaactcctccgctttttacgtgttgaaaaagactgatcagcacgaaacggggagct
aagctaccgcttcagcacttgagagcagtatcttccatccaccgccgttcgtcaggggcaaggctcag
atcaacgagcgcctccatttacacggagcggggatcccaacgtccacactgtgctgtcacccacgcgac
gcaaccctacccagccaccaacaccatcaggtccctcagaagaactcgtccaacagccggtaaaacgcc
agcttttcctccgataccgccccatcccacccgcgcccgtactcccgcaggaacgccgcggaacactcc
ggcccgaaccacgggtcctcctcgtgggccagctcgcgcagcaccagcgcgagatcggagtgccggtcc
gcacggccgacccgccccacgtcgatcagcccggtcacctcgcaggtacgagggtcgagcagcacgttg
tccgggcacaggtcaccgtggcaaaccgccagatcctcgtccgcaggccgagtccgctccagctcggcg
agaagccgctcccccgaccacccccttccgctcctcgtccagatcctccaagtcgacgctcccttcagcg
acagcacgggccgcctgcggcaccgtcaccgcgagactgcgatcgaacggacaccgctcccagtccagc
gcgtgcagcgaacgagcgagccccgcgagcgccaccgccacgtccagccgctgctcccgcggccaccgc
gcactggccggacgccccggaaccgcttcggtgaccaaccaggcgaccctctcgtccccaccaccctcc
acaacacgaggtacgggaatccccacctccgccaaccacaccagccgctcagcctcacccaacaagccc
accccggcccccagagctgccaccttgacaaacaactcccgcccaccaccccgaagccgataaacacca
gcccccgaggccccatcctccacaacaacccactcacaaccgggataccgaccccgcagtgcacgcaac
gcatcgtccatgcttcgaaattcttcagcaccggggagggcggagtggccatcctgcaaatggaaacgg
cgacgcagggttagatgctgcttgagacagcgacagaggagccaaaagccttcgtcgacacaatgcggg
cgttgcaagtcaaatctgcaagcacgctgcctgatccgccgggcttgctcgtcgactcacctggccatt
ttaagatgttgagtgacttctcttgtaaaaaagtaaagaacataggccccctggccggtttatcaggag
ggcaccgctccaggggctgcatgcgaactgcttgcattggcgcctagcctttgtgggccaggggcttc
cggataagggttgcaagtgctcaaataccccatcaaacatcatcctggtttggctgcgctccttctggc
gcgcccggcatgcaagcttgatgggatcttaagctagctgagtggttatgtatagcggcagaatagtcg
cgtatgtataagtgctcgtttgtcgctgaaagtggaggtcaccgttcggggtcgcgggcttttataccg
gatgggtgccgccagcgggccgtatggcgccttctggacgccgcgcgccccatcgcggcccttccagag
cttcccccgccctcatagcccgccaaatcagtcctgtagcttcatacaaacatacgcaccaatcatgtc
aagcctcagcgagctccccgctcgagcacacacctgcccgtctgcctgacaggaagtgaacgcatgtcg
agggaggcctcaccaatcgtcacacgagccctcgtcagaaacacgtctccgccacgctctccctctcac
ggccgaccccgcagccctttgccctttcctaggccaccgacaggacccaggcgctctcagcatgcctc
aacaacccgtactcgtgccagcggtgcccttgtgctggtgatcgcttggaagcgcatgcgaagacgaag
gggcggagcaggcggcctggctgttcgaagggctcgccgccagttcgggtgcctttctccacgcgcgcc
tccacacctaccgatgcgtgaaggcaggcaaatgctcatgtttgcccgaactcggagtccttaaaaagc
cgcttcttgtcgtcgttccgagacatgttagcagatcgcagtgccacctttcctgacgcgctcggcccc
atattcggacgcaattgtcatttgtagcacaattggagcaaatctggcgaggcagtaggcttttaagtt
gcaaggcgagagagcaaagtgggacgcggcgtgattattggtatttacgcgacggcccggcgcgttagc
ggcccttcccccaggccagggacgattatgtatcaatattgttgcgttcgggcactcgtgcgagggctc
ctgcgggctggggaggggatctgggaattggaggtacgaccgagatggcttgctcgggggggaggtttc
ctcgccgagcaagccagggttaggtgttgcgctcttgactcgttgtgcattctaggaccccactgctac
tcacaacaagcccatatgatcgataagcttgatatcgaattcctgcagcccgggggatccactagttct
agatggcagcagctggaccgcctgtaccatggagaagagctttacttgccgggatggccgatttcgctg
attgatacgggatcggagctcggaggctttcgcgctaggggctaggcgaagggcagtggtgaccagggt
cggtgtggggtcggcccacggtcaattagccacaggaggatcaggggaggtaggcacgtcgacttggt
ttgcgaccccgcagttttggcggacgtgctgttgtagatgttagcgtgtgcgtgagccagtggccaacg
tgccacacccattgagaagaccaaccaacttactggcaatatctgccaatgccatactgcatgtaatgg
ccaggccatgtgagagtttgccgtgcctgcgcgcgccccgggggcgcagtttagctgaccagccgtggg
atgatgcacgcatttgcaaggacagggtaatcacagcagcaacatggtgggcttaggacagctgtgggt
cagtggacggacggcaggggagggacggcgcagctcgggagacaggggggagacagcgtgactgtgcaat

gcggccgccaccgcggtggagctccaattcgccctatagtgagtcgtattacaattcactggccgtcgt
tttacaacgtcgtgactgggaaaaccctggcgttacccaacttaatcgccttgcagcacatccccctttt
cgccagctggcgtaatagcgaagaggcccgcaccgatcgcccttcccaacagttgcgcagcctgaatgg
cgaatggaaattgtaagcgttaatattttgttaaaattcgcgttaaattttgttaaatcagctcattt
tttaaccaataggccgaaatcggcaaaatcccttataaatcaaaagaatagaccgagatagggttgagt
gttgttccagtttggaacaagagtccactattaaagaacgtggactccaacgtcaaagggcgaaaaacc
gtctatcagggcgatggcccactacgtgaaccatcaccctaatcaagtttttggggtcgaggtgccgt
aaagcactaaatcggaaccctaaagggagcccccgatttagagcttgacggggaaagccggcgaacgtg
gcgagaaaggaagggaagaaagcgaaaggagcgggcgctagggcgctggcaagtgtagcggtcacgctg
cgcgtaaccaccacacccgccgcgcttaatgcgccgctacagggcg

Figure 5: Genewiz synthesized sequence of *Auxenochlorella protothecoides* (formerly known as *Chlorella protothecoides* (CP)) promoter pair inserted into pSL18 at the NheI and NotI restriction sites in pSL18. Blue highlight = CP Psad Promoter sequence; Red highlight = CP Psad Terminator sequence; Yellow, Pink and Green = various restriction enzyme sites for cloning

tccttctggcgcgcccggcatgcaagcttgatgggatcttaagctagcAATTCCAGCCTGCCCTGTTCA
TGCCTCTGGCATTGAACTTTATTCTCATCGATGCTGTTGTAGCTTTACAGTGACCAAGATGGACAGGTG
TACAGTGGGCGTTGCATGAAAGGGTGGAAGTGATGAGGAAATCAAAATTGGGTACTAGGTCAGGGAAGG
TGGAGGTCGTGGGCAAAAGGGGGTGCTCCGGATGCATGGTGGGTTTGAGAGGAGTAAAGTATAATCTGA
CAGCCGATGGGGCAACAATTGCGTGTGAGACAGTGAAGTCAGAGGACTCCATCGCGCATGAAGCAAGAG
AGCCTTCCGAGGCCCTGGAACACTCGAACAGCATCCTGCAGGATCCTTATATTTCTCTGACTAGGCAGC
ATTGGTTATTTGCCGTGCCAGGGCCCACTCTAGCCCAGACTCCCCAGACTCCAGATCGAGGTCAGCAGC
GTCCTCGGAGTCCCTCTCCTCCGCAGCAGTAAGGCAGTCAACCGACTGCAGCTTGCGGAGACTGTTGAG
GCCTGCACATGTGTTGGTTGCCAGAGGGAGTGACAAGGGGTGAAAGGTAAGGATGTACAAGGCACGCAC
CATGGCAGGCGATTGCAGCCATATTTCAAGGGCATCTTGAAACTGCTGTCGTGGGAAACTCTCACTGGC
CCCGCTCTCCCAGCTCTACTGTAAAGTTGCCAGATCCCACCATCAGCTCACCATTCTGGTCTGAGGTCG
CCACCACCGACGTCTGCGCGGCATTGAAGGCGTCTGCAGCCGCCTGGGCGGCAGCGACGTTGCGCATGG
TGTCATCCAAATCGCTCTGCTCGATCACCTCCCCCTCATCATCATACATGGGAATGTCCTCGTCCTGTA
TTATGGGGTGCATTTATGGCCCGTAGGGCAGACAGCACGCAAGGAAAGGACAAAGACGCGACGGCAGCG
AGTCACCACAACCTGCACCAATGTCAACAGTGCTAAGATTCTGGTCCTCGCTCTCCATCCACCACTCCG
GGGTCGCATATCCGACGCCCCAGGAATCGTTCCCAAATCCACTCTCCATCGAGTTCCCGGCTGAGGCAG
GGCGGTACCAGATCTGTATACCATATGTTAATTAAGAGCTCTCTAGAGTCGACCTGCAGGGGGGCAGTG
AGTTATCAGGAACAGGGTGGAGCCTGTAAGGGGATGCATCTCAGCGGCTCAAGCCGAGGACATCTGAGG
CCGGAACCTCCACAGTTTCTACTCTCTCCCTGTATGATGAGCAGCTGCACGACCCTTTGTCGAGAAGAA
TATGTGTACCGAGGATGTGGTGCTGTCAACAACACAACACTGCGAGGAAGGGGGCCAGGGGCATGTTTG
TACAGGGGCGTGCACCGCCAGGCATCTGCATGGTGCCGAGTCACCAGCATGTCAGAGGCAGCTGCTATT
GGGGAAAGGGCTTGGGATACCCGGGGTTTGTTGTGCATGCTTATCGTGCATGTGGCACAGGGCAGCAGC
AAGCCGGGTCTGGTCAGGGAAGGCTTCCTGGGGTCCTTGAGATGGGCTTTGCCACATCCAGTCCCCTGC
ACGCCTCGACCCCGGGCCCTGGGGCCCGTCCACCTACCCCTCCTCCCCAAACTCGGTTGTGAACTTCTT
GTACGTCTCCAGGTCGTCGAATTCCAGCCTGCCCTGTTCATGCCTCTGGCATTGAACTTTATTCTCATC
GATGCTGTTGTAGCTTTACAGTGACCAAGATGGACAGGTGTACAGTGGGCGTTGCATGAAAGGGTGGAA
GTGATGAGGAAATCAAAATTGGGTACTAGGTCAGGGAAGGTGGAGGTCGTGGGCAAAAGGGGGTGCTCC
GGATGCATGGTGGGTTTGAGAGGAGTAAAGTATAATCTGACAGCCGATGGGGCAACAATTGCGTGTGAG
ACAGTGAAGTCAGAGGACTCCATCGCGCATGAAGCAAGAGAGCCTTCCGAGGCCCTGGAACACTCGAAC
AGCATCCTGCAGGATCCTTATATTTCTCTGACTAGGCAGCATTGGTTATTTGCCGTGCCAGGGCCCACT
CTAGCCCAGACTCCCCAGACTCCAGATCGAGGTCAGCAGCGTCCTCGGAGTCCCTCTCCTCCGCAGCAG
TAAGGCAGTCAACCGACTGCAGCTTGCGGAGACTGTTGAGGCCTGCACATGTGTTGGTTGCCAGAGGGA
GTGACAAGGGGTGAAAGGTAAGGATGTACAAGGCACGCACCATGGCAGGCGATTGCAGCCATATTTCAA
GGGCATCTTGAAACTGCTGTCGTGGGAAACTCTCACTGGCCCCGCTCTCCCAGCTCTACTGTAAAGTTG
CCAGATCCCACCATCAGCTCACCATTCTGGTCTGAGGTCGCCACCACCGACGTCTGCGCGGCATTGAAG
GCGTCTGCAGCCGCCTGGGCGGCAGCGACGTTGCGCATGGTGTCATCCAAATCGCTCTGCTCGATCACC

TCCCCCTCATCATCATACATGGGAATGTCCTCGTCCTGTATTATGGGGTGCATTTATGGCCCGTAGGGC
AGACAGCACGCAAGGAAAGGACAAAGACGCGACGGCAGCGAGTCACCACAACCTGCACCAATGTCAACA
GTGCTAAGATTCTGGTCCTCGCTCTCCATCCACCACTCCGGGGTCGCATATCCGACGCCCCAGGAATCG
TTCCCAAATCCACTCTCCATCGAGTTCCCGGCTGAGGCAGGGCGAATGCTGATATCACTAGTCTCGAGG
GCCGGCCAAGCTTGTCGACCTGCAGGGGGGCAGTGAGTTATCAGGAACAGGGTGGAGCCTGTAAGGGGA
TGCATCTCAGCGGCTCAAGCCGAGGACATCTGAGGCCGGAACCTCCACAGTTTCTACTCTCTCCCTGTA
TGATGAGCAGCTGCACGACCCTTTGTCGAGAAGAATATGTGTACCGAGGATGTGGTGCTGTCAACAACA
CAACACTGCGAGGAAGGGGGCCAGGGGCATGTTTGTACAGGGGCGTGCACCGCCAGGCATCTGCATGGT
GCCGAGTCACCAGCATGTCAGAGGCAGCTGCTATTGGGGAAAGGGCTTGGGATACCCGGGGTTTGTTGT
GCATGCTTATCGTGCATGTGGCACAGGGCAGCAGCAAGCCGGGTCTGGTCAGGGAAGGCTTCCTGGGGT
CCTTGAGATGGGCTTTGCCACATCCAGTCCCCTGCACGCCTCGACCCCGGGCCCTGGGGCCCGTCCACC
TACCCCTCCTCCCCAAACTCGGTTGTGAACTTCTTGTACGTCTCCAGGTCGTCgcggccgccaccgcgg
tggagctccaattcgccctatagtgagtcgtat

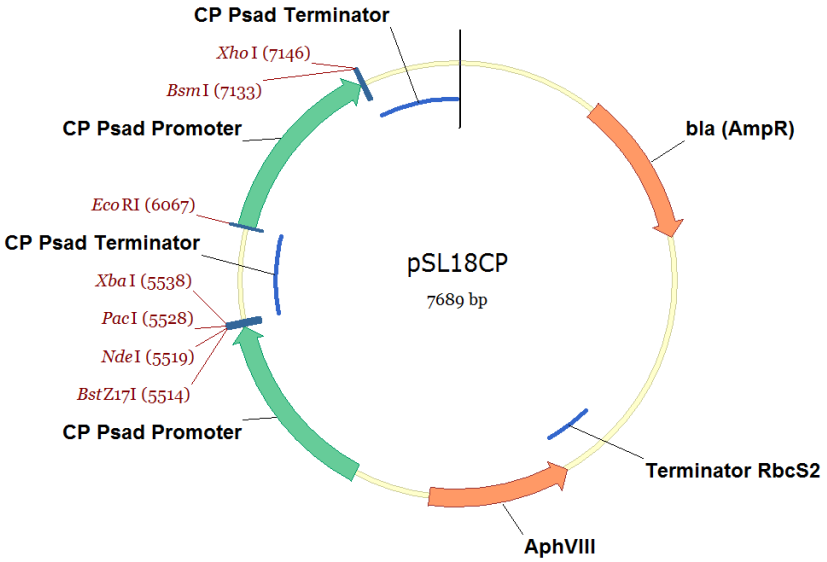


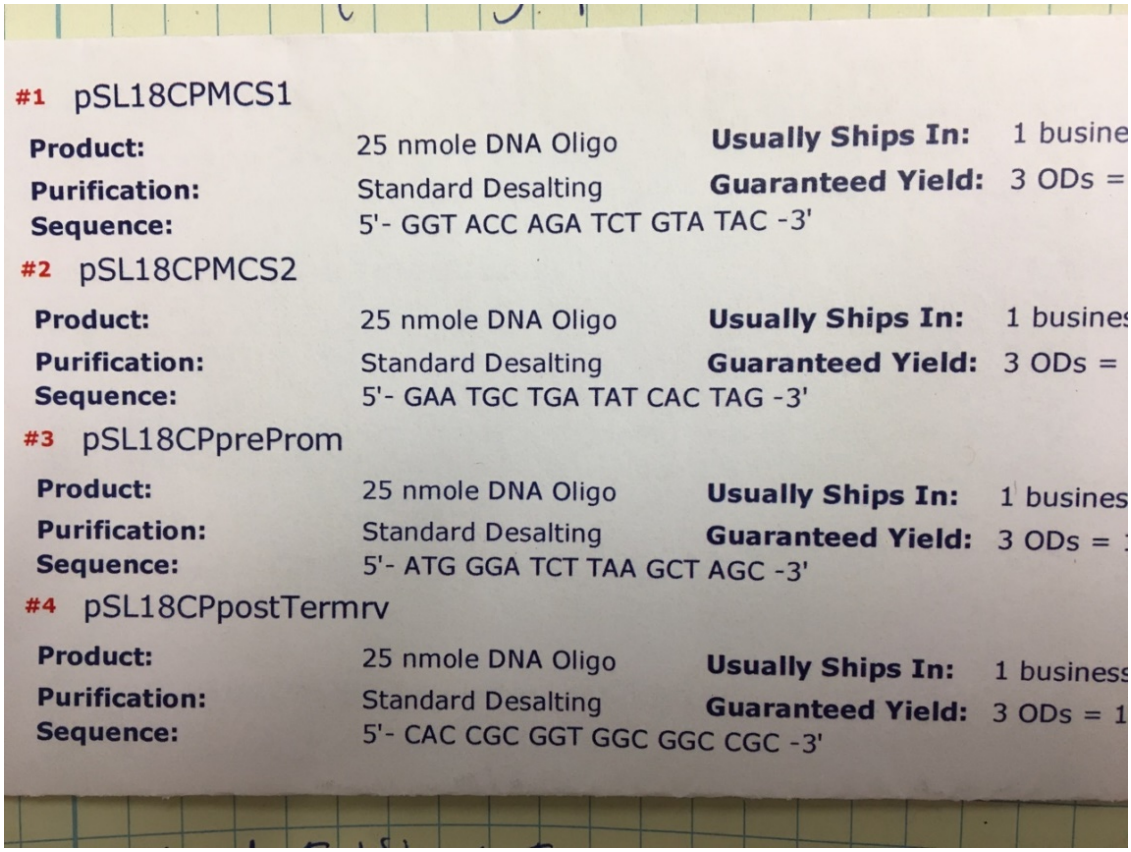
Figure 6: Map of pSL18CP plasmid showing key restriction enzyme sites engineered into the CP Psad promoter/terminator insert. CP Psad Promoter = *Auxenochlorella protothecoides* promoter; CP Psad Terminator = *Auxenochlorella protothecoides* terminator; bla = Ampicillin resistance gene; AphVIII = Paromycin resistance gene.

Figure 6A: Sequence of pSL18CP plasmid (map shown in Figure 6).

ggccgccaccgcggtggagctccaattcgccctatagtgagtcgtattacgcgcgctcactggccgtcg
ttttacaacgtcgtgactgggaaaaccctggcgttacccaacttaatcgccttgcagcacatcccccctt
tcgccagctggcgtaatagcgaagaggcccgcaccgatcgcccttcccaacagttgcgcagcctgaatg
gcgaatgggacgcgccctgtagcggcgcattaagcgcggcgggtgtggtggttacgcgcagcgtgaccg
ctacacttgccagcgccctagcgcccgctcctttcgctttcttcccttcctttctcgccacgttcgccg
gctttccccgtcaagctctaaatcgggggctccctttagggttccgatttagtgctttacggcacctcg
accccaaaaaacttgattagggtgatggttcacgtagtgggccatcgccctgatagacggtttttcgcc
ctttgacgttggagtccacgttctttaatagtggactcttgttccaaactggaacaacactcaacccta
tctcggtctattcttttgatttataagggattttgccgatttcggcctattggttaaaaaatgagctga
tttaacaaaaatttaacgcgaattttaacaaaatattaacgcttacaatttagcccggggtggcacttt
tcggggaaatgtgcgcggaacccctatttgtttatttttctaaatacattcaaatatgtatccgctcat
gagacaataaccctgataaatgcttcaataatattgaaaaaggaagagtatgagtattcaacatttccg
tgtcgcccttattccctttttgcggcattttgccttcctgtttttgctcacccagaaacgctggtgaa
agtaaaagatgctgaagatcagttgggtgcacgagtgggttacatcgaactggatctcaacagcggtaa

gatccttgagagttttcgccccgaagaacgttttccaatgatgagcacttttaaagttctgctatgtgg
cgcggtattatcccgtattgacgccgggcaagagcaactcggtcgccgcatacactattctcagaatga
cttggttgagtactcaccagtcacagaaaagcatcttacggatggcatgacagtaagagaattatgcag
tgctgccataaccatgagtgataacactgcggccaacttacttctgacaacgatcggaggaccgaagga
gctaaccgcttttttgcacaacatgggggatcatgtaactcgccttgatcgttgggaaccggagctgaa
tgaagccataccaaacgacgagcgtgacaccacgatgcctgtagcaatggcaacaacgttgcgcaaact
attaactggcgaactacttactctagcttcccggcaacaattaatagactggatggaggcggataaagt
tgcaggaccacttctgcgctcggcccttccggctggctggtttattgctgataaatctggagccggtga
gcgtgggtctcgcggtatcattgcagcactggggccagatggtaagccctcccgtatcgtagttatcta
cacgacggggagtcaggcaactatggatgaacgaaatagacagatcgctgagataggtgcctcactgat
taagcattggtaactgtcagaccaagtttactcatatatactttagattgatttaaaacttcatttta
atttaaaaggatctaggtgaagatccttttgataatctcatgaccaaaatcccttaacgtgagttttc
gttccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatccttttttctgcgcgt
aatctgctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgccggatcaagagctacc
aactcttttccgaaggtaactggcttcagcagagcgcagataccaaatactgtccttctagtgtagcc
gtagttaggccaccacttcaagaactctgtagcaccgcctacatacctcgctctgctaatcctgttacc
agtggctgctgccagtggcgataagtcgtgtcttaccgggttggactcaagacgatagttaccggataa
ggcgcagcggtcgggctgaacgggggttcgtgcacacagcccagcttggagcgaacgacctacaccga
actgagatacctacagcgtgagctatgagaaagcgccacgcttcccgaagggagaaaggcggacaggta
tccggtaagcggcagggtcggaacaggagagcgcacgagggagcttccagggggaaacgcctggtatct
ttatagtcctgtcgggtttcgccacctctgacttgagcgtcgattttgtgatgctcgtcaggggggcg
gagcctatggaaaaacgccagcaacgcggccttttacggttcctggccttttgctggccttttgctca
catgttctttcctgcgttatcccctgattctgtggataaccgtattaccgcctttgagtgagctgatac
cgctcgccgcagccgaacgaccgagcgcagcgagtcagtgagcgaggaagcggaagagcgcccaatacg
caaaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggtttcccgactggaa
agcgggcagtgagcgcaacgcaattaatgtgagttagctcactcattaggcaccccaggctttacactt
tatgcttccggctcgtatgttgtgtggaattgtgagcggataacaatttcacacaggaaacagctatga
ccatgattacgccaagcgcgcaattaaccctcactaaagggaacaaaagctgggtacccgcttcaaata
cgcccagcccgcccatggagaaagaggccaaaatcaacggaggatcgttacaaccaacaaaattgcaaa
actcctccgctttttacgtgttgaaaaagactgatcagcacgaaacggggagctaagctaccgcttcag
cacttgagagcagtatcttccatccaccgccgttcgtcagggggcaaggctcagatcaacgagcgcctc
catttacacggagcggggatcccaacgtccacactgtgctgtcacccacgcgacgcaaccctacccagc
caccaacaccatcaggtccctcagaagaactcgtccaacagccggtaaaacgccagcttttcctccgat
accgccccatcccacccgcgcccgtactcccgcaggaacgccgcggaacactccggcccgaaccacggg
tcctcctcgtgggccagctcgcgcagcaccagcgcgagatcggagtgccggtccgcacggccgacccgc
cccacgtcgatcagcccggtcacctcgcaggtacgagggtcgagcagcacgttgtccgggcacaggtga
ccgtggcaaaccgccagatcctcgtccgcaggccgagtccgctccagctcggcgagaagccgctccccc
gaccacccttccgctcctcgtccagatcctccaagtcgacgctcccttcagcgacagcacgggccgcc
tgcggcaccgtcaccgcgagactgcgatcgaacggacaccgctcccagtccagcgcgtgcagcgaacga
gcgagccccgcgagcgccaccgccacgtccagccgctgctcccgcggccaccgcgcactggccggacgc
cccggaaccgcttcggtgaccaaccaggcgaccctctcgtccccaccaccctccacaacacgaggtacg
ggaatccccacctccgccaaccacaccagccgctcagcctcacccaacaagcccacccggcccccaga
gctgccaccttgacaaacaactcccgcccaccaccccgaagccgataaacaccagcccccgaggcccca
tcctccacaacaacccactcacaaccgggataccgaccccgcagtgcacgcaacgcatcgtccatgctt
cgaaattcttcagcaccggggagggcggagtggccatcctgcaaatggaaacggcgacgcagggttaga
tgctgcttgagacagcgacagaggagccaaaagccttcgtcgacacaatgcgggcgttgcaagtcaaat
ctgcaagcacgctgcctgatccgccgggcttgctcgtcgactcacctggccattttaagatgttgagtg
acttctcttgtaaaaaagtaaagaacataggccccctggccggtttatcaggagggcaccgctccaggg
gctgcatgcgaactgcttgcattggcgcctagcctttgtgggccaggggcttccggataagggttgca
agtgctcaaataccccatcaaacatcatcctggtttggctgcgctccttctggcgcgcccggcatgcaa
gcttgatgggatcttaagctagcaattccagcctgccctgttcatgcctctggcattgaactttattct

catcgatgctgttgtagctttacagtgaccaagatggacaggtgtacagtgggcgttgcatgaaagggt
ggaagtgatgaggaaatcaaaattgggtactaggtcagggaaggtggaggtcgtgggcaaaaggggtg
ctccggatgcatggtgggtttgagaggagtaaagtataatctgacagccgatggggcaacaattgcgtg
tgagacagtgaagtcagaggactccatcgcgcatgaagcaagagagccttccgaggccctggaacactc
gaacagcatcctgcaggatccttatatttctctgactaggcagcattggttatttgccgtgccagggcc
cactctagcccagactccccagactccagatcgaggtcagcagcgtcctcggagtccctctcctccgca
gcagtaaggcagtcaaccgactgcagcttgcggagactgttgaggcctgcacatgtgttggttgccaga
gggagtgacaaggggtgaaaggtaaggatgtacaaggcacgcaccatggcaggcgattgcagccatatt
tcaagggcatcttgaaactgctgtcgtgggaaactctcactggccccgctctcccagctctactgtaaa
gttgccagatcccaccatcagctcaccattctggtctgaggtcgccaccaccgacgtctgcgcggcatt
gaaggcgtctgcagccgcctgggcggcagcgacgttgcgcatggtgtcatccaaatcgctctgctcgat
cacctccccctcatcatcatacatgggaatgtcctcgtcctgtattatggggtgcatttatggcccgta
gggcagacagcacgcaaggaaaggacaaagacgcgacggcagcgagtcaccacaacctgcaccaatgtc
aacagtgctaagattctggtcctcgctctccatccaccactccggggtcgcatatccgacgccccagga
atcgttcccaaatccactctccatcgagttcccggctgaggcagggcggtaccagatctgtataccata
tgttaattaagagctctctagagtcgacctgcaggggggcagtgagttatcaggaacagggtggagcct
gtaaggggatgcatctcagcggctcaagccgaggacatctgaggccggaacctccacagtttctactct
ctccctgtatgatgagcagctgcacgacccttgtcgagaagaatatgtgtaccgaggatgtggtgctg
tcaacaacacaacactgcgaggaaggggccaggggcatgtttgtacaggggcgtgcaccgccaggcat
ctgcatggtgccgagtcaccagcatgtcagaggcagctgctattggggaaagggcttgggatacccggg
gtttgttgtgcatgcttatcgtgcatgtggcacagggcagcagcaagccgggtctggtcagggaaggct
tcctggggtccttgagatgggctttgccacatccagtcccctgcacgcctcgaccccgggccctggggc
ccgtccacctacccctcctccccaaactcggttgtgaacttcttgtacgtctccaggtcgtcgaattcc
agcctgccctgttcatgcctctggcattgaactttattctcatcgatgctgttgtagctttacagtgac
caagatggacaggtgtacagtgggcgttgcatgaaagggtggaagtgatgaggaaatcaaaattgggta
ctaggtcagggaaggtggaggtcgtgggcaaaaggggtgctccggatgcatggtgggtttgagaggag
taaagtataatctgacagccgatggggcaacaattgcgtgtgagacagtgaagtcagaggactccatcg
cgcatgaagcaagagagccttccgaggccctggaacactcgaacagcatcctgcaggatccttatattt
ctctgactaggcagcattggttatttgccgtgccagggcccactctagcccagactccccagactccag
atcgaggtcagcagcgtcctcggagtccctctcctccgcagcagtaaggcagtcaaccgactgcagctt
gcggagactgttgaggcctgcacatgtgttggttgccagagggagtgacaaggggtgaaaggtaaggat
gtacaaggcacgcaccatggcaggcgattgcagccatatttcaagggcatcttgaaactgctgtcgtgg
gaaactctcactggccccgctctcccagctctactgtaaagttgccagatcccaccatcagctcaccat
tctggtctgaggtcgccaccaccgacgtctgcgcggcattgaaggcgtctgcagccgcctgggcggcag
cgacgttgcgcatggtgtcatccaaatcgctctgctcgatcacctccccctcatcatcatacatgggaa
tgtcctcgtcctgtattatggggtgcatttatggcccgtagggcagacagcacgcaaggaaaggacaaa
gacgcgacggcagcgagtcaccacaacctgcaccaatgtcaacagtgctaagattctggtcctcgctct
ccatccaccactccggggtcgcatatccgacgccccaggaatcgttcccaaatccactctccatcgagt
tcccggctgaggcagggcgaatgctgatatcactagtctcgagggccggccaagcttgtcgacctgcag
gggggcagtgagttatcaggaacagggtggagcctgtaaggggatgcatctcagcggctcaagccgagg
acatctgaggccggaacctccacagtttctactctctccctgtatgatgagcagctgcacgacccttttg
tcgagaagaatatgtgtaccgaggatgtggtgctgtcaacaacacaacactgcgaggaaggggccaggg
ggcatgtttgtacaggggcgtgcaccgccaggcatctgcatggtgccgagtcaccagcatgtcagaggc
agctgctattggggaaagggcttgggatacccggggtttgttgtgcatgcttatcgtgcatgtggcaca
gggcagcagcaagccgggtctggtcagggaaggcttcctggggtccttgagatgggctttgccacatcc
agtcccctgcacgcctcgaccccgggccctggggcccgtccacctacccctcctccccaaactcggttg
tgaacttcttgtacgtctccaggtcgtcgc



#1 pSL18CPMCS1

Product:	25 nmole DNA Oligo	**Usually Ships In:**	1 busine
Purification:	Standard Desalting	**Guaranteed Yield:**	3 ODs =
Sequence:	5'- GGT ACC AGA TCT GTA TAC -3'		

#2 pSL18CPMCS2

Product:	25 nmole DNA Oligo	**Usually Ships In:**	1 busines
Purification:	Standard Desalting	**Guaranteed Yield:**	3 ODs =
Sequence:	5'- GAA TGC TGA TAT CAC TAG -3'		

#3 pSL18CPpreProm

Product:	25 nmole DNA Oligo	**Usually Ships In:**	1 busines
Purification:	Standard Desalting	**Guaranteed Yield:**	3 ODs =
Sequence:	5'- ATG GGA TCT TAA GCT AGC -3'		

#4 pSL18CPpostTermrv

Product:	25 nmole DNA Oligo	**Usually Ships In:**	1 business
Purification:	Standard Desalting	**Guaranteed Yield:**	3 ODs = 1
Sequence:	5'- CAC CGC GGT GGC GGC CGC -3'		

Figure 7: List of primers used by Genewiz to sequence confirm the pSL18CP plasmid vector.

The pSL18CP vector was further engineered by researchers at LANL to include an *A. protothecoides* codon optimized copy of AphVIII (paramomycin resistance gene) (sequence in Figure 8) at the XhoI and BsmI restriction sites. This became the intermediate vector pCpAphVIII shown in Figure 9 with sequence immediately following in Figure 9A. The PCR gel and primers used to confirm AphVIII insertion in pCpAphVIII are shown in Figure 10. Note that the *A. protothecoides* codon optimization of AphVIII did not alter the nucleotide sequence from the commercial paramomycin-containing vectors.

Figure 8: Sequence of *A. protothecoides* codon optimized copy of AphVIII (paramomycin resistance gene) inserted into at the XhoI and BsmI restriction sites of pSL18CP to create pCpAphVIII.

Atggacgatgcgttgcgtgcactgcggggtcggtatcccggttgtgagtgggttgttgtggaggatggg
gcctcgggggctggtgtttatcggcttcggggtggtgggcgggagttgtttgtcaaggtggcagctctg
ggggccggggtgggcttgttgggtgaggctgagcggctggtgtggttggcggaggtggggattcccgta
cctcgtgttgtggagggtggtggggacgagagggtcgcctggttggtcaccgaagcggttccggggcgt
ccggccagtgcgcggtggccgcgggagcagcggctggacgtggcggtggcgctcgcggggctcgctcgt
tcgctgcacgcgctggactgggagcggtgtccgttcgatcgcagtctcgcggtgacggtgccgcaggcg
gcccgtgctgtcgctgaagggagcgtcgacttggaggatctggacgaggagcggaaggggtggtcgggg
gagcggcttctcgccgagctggagcggactcggcctgcggacgaggatctggcggtttgccacggtcac
ctgtgcccggacaacgtgctgctcgaccctcgtacctgcgaggtgaccgggctgatcgacgtggggcgg
gtcggccgtgcggaccggcactccgatctcgcgctggtgctgcgcgagctggcccacgaggaggacccg
tggttcgggccggagtgttccgcggcgttcctgcgggagtacgggcgcgggtgggatggggcggtatcg
gaggaaaagctggcgttttaccggctgttggacgagttcttctga

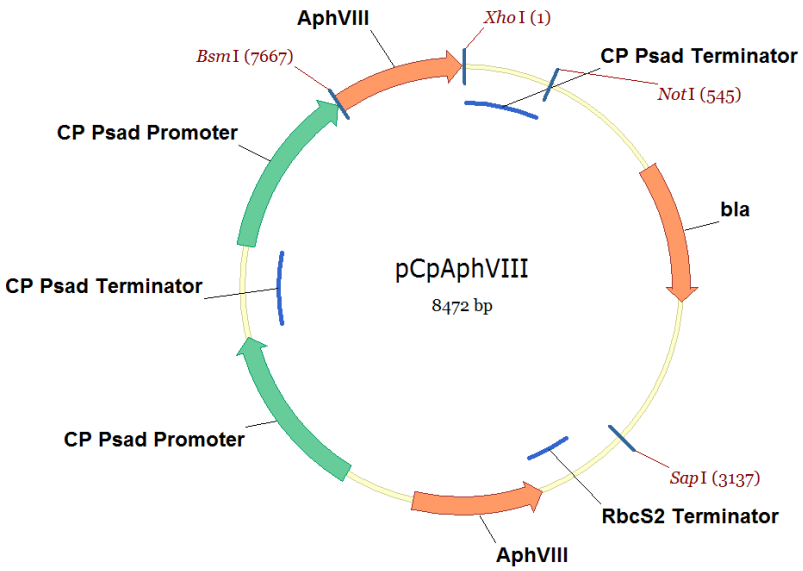


Figure 9: Map of pCpAphVIII plasmid showing key restriction enzyme sites. CP Psad Promoter = *Auxenochlorella protothecoides* promoter; CP Psad Terminator = *Auxenochlorella protothecoides* terminator; bla = Ampicillin resistance gene; AphVIII = Paromycin resistance gene.

Figure 9A: Sequence of pCpAphVIII plasmid (map shown in Figure 9).

```
tcgagggccggccaagcttgtcgacctgcaggggggcagtgagttatcaggaacagggtggagcctgta
aggggatgcatctcagcggctcaagccgaggacatctgaggccggaacctccacagtttctactctctc
cctgtatgatgagcagctgcacgacccttgtcgagaagaatatgtgtaccgaggatgtggtgctgtca
acaacacaacactgcgaggaaggggccaggggcatgtttgtacaggggcgtgcaccgccaggcatctg
catggtgccgagtcaccagcatgtcagaggcagctgctattggggaaagggcttgggatacccggggtt
tgttgtgcatgcttatcgtgcatgtggcacagggcagcagcaagccgggtctggtcagggaaggcttcc
tggggtccttgagatgggctttgccacatccagtcccctgcacgcctcgaccccgggccctggggcccg
tccacctacccctcctccccaaactcggttgtgaacttcttgtacgtctccaggtcgtcgcggccgcca
ccgcggtggagctccaattcgccctatagtgagtcgtattacaattcactggccgtcgttttacaacgt
cgtgactgggaaaaccctggcgttacccaacttaatcgccttgcagcacatccccctttcgccagctgg
cgtaatagcgaagaggcccgcaccgatcgcccttcccaacagttgcgcagcctgaatggcgaatggaaa
ttgtaagcgttaatattttgttaaaattcgcgttaaatttttgttaaatcagctcatttttaaccaat
aggccgaaatcggcaaaatcccttataaatcaaaagaatagaccgagatagggttgagtgttgttccag
tttggaacaagagtccactattaaagaacgtggactccaacgtcaaagggcgaaaaaccgtctatcagg
gcgatggcccactacgtgaaccatcaccctaatcaagtttttggggtcgaggtgccgtaaagcactaa
atcggaaccctaaagggagcccccgatttagagcttgacggggaaagccggcgaacgtggcgagaaagg
aagggaagaaagcgaaaggagcgggcgctagggcgctggcaagtgtagcggtcacgctgcgcgtaacca
ccacacccgccgcgcttaatgcgccgctacagggcgcgtcaggtggcacttttcggggaaatgtgcgcg
gaacccctatttgtttatttttctaaatacattcaaatatgtatccgctcatgagacaataaccctgat
aaatgcttcaataatattgaaaaaggaagagtatgagtattcaacatttccgtgtcgcccttattccct
tttttgcggcattttgccttcctgtttttgctcacccagaaacgctggtgaaagtaaaagatgctgaag
atcagttgggtgcacgagtgggttacatcgaactggatctcaacagcggtaagatccttgagagtttc
gccccgaagaacgttttccaatgatgagcacttttaaagttctgctatgtggcgcggtattatcccgta
ttgacgccgggcaagagcaactcggtcgccgcatacactattctcagaatgacttggttgagtactcac
cagtcacagaaaagcatcttacggatggcatgacagtaagagaattatgcagtgctgccataaccatga
gtgataacactgcggccaacttacttctgacaacgatcggaggaccgaaggagctaaccgcttttttgc
acaacatgggggatcatgtaactcgccttgatcgttgggaaccggagctgaatgaagccataccaaacg
acgagcgtgacaccacgatgcctgtagcaatggcaacaacgttgcgcaaactattaactggcgaactac
```

ttactctagcttcccggcaacaattaatagactggatggaggcggataaagttgcaggaccacttctgc
gctcggcccttccggctggctggtttattgctgataaatctggagccggtgagcgtgggtctcgcggta
tcattgcagcactggggccagatggtaagccctcccgtatcgtagttatctacacgacggggagtcagg
caactatggatgaacgaaatagacagatcgctgagataggtgcctcactgattaagcattggtaactgt
cagaccaagtttactcatatatactttagattgatttaaaacttcatttttaatttaaaaggatctagg
tgaagatccttttttgataatctcatgaccaaaatcccttaacgtgagttttcgttccactgagcgtcag
accccgtagaaaagatcaaaggatcttcttgagatcctttttttctgcgcgtaatctgctgcttgcaaa
caaaaaaaccaccgctaccagcggtggtttgtttgccggatcaagagctaccaactctttttccgaagg
taactggcttcagcagagcgcagataccaaatactgttcttctagtgtagccgtagttaggccaccact
tcaagaactctgtagcaccgcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtg
gcgataagtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcgcagcggtcgggct
gaacgggggttcgtgcacacagcccagcttggagcgaacgacctacaccgaactgagatacctacagc
gtgagctatgagaaagcgccacgcttcccgaagggagaaaggcggacaggtatccggtaagcggcaggg
tcggaacaggagagcgcacgagggagcttccagggggaaacgcctggtatctttatagtcctgtcgggt
ttcgccacctctgacttgagcgtcgattttgtgatgctcgtcaggggggcggagcctatggaaaaacg
ccagcaacgcggccttttacggttcctggccttttgctggccttttgctcacatgttctttcctgcgt
tatcccctgattctgtggataaccgtattaccgcctttgagtgagctgataccgctcgccgcagccgaa
cgaccgagcgcagcgagtcagtgagcgaggaagcggaagagcgcccaatacgcaaaccgcctctccccg
cgcgttggccgattcattaatgcagctggcacgacaggtttcccgactggaaagcgggcagtgagcgca
acgcaattaatgtgagttagctcactcattaggcaccccaggctttacactttatgcttccggctcgta
tgttgtgtggaattgtgagcggataacaatttcacacaggaaacagctatgaccatgattacgccaagc
tcgaaattaaccctcactaaagggaacaaaagctggtacccgcttcaaatacgcccagcccgcccatgg
agaaagaggccaaaatcaacggaggatcgttacaaccaacaaaattgcaaaactcctccgcttttttacg
tgttgaaaaagactgatcagcacgaaacggggagctaagctaccgcttcagcacttgagagcagtatct
tccatccaccgccgttcgtcaggggcaaggctcagatcaacgagcgcctccatttacacggagcgggg
atcccaacgtccacactgtgctgtcacccacgcgacgcaaccctacccagccaccaacaccatcaggtc
cctcagaagaactcgtccaacagccggtaaaacgccagcttttcctccgataccgccccatcccacccg
cgcccgtactcccgcaggaacgccgcggaacactccggcccgaaccacgggtcctcctcgtgggccagc
tcgcgcagcaccagcgcgagatcggagtgccggtccgcacggccgacccgccccacgtcgatcagcccg
gtcacctcgcaggtacgagggtcgagcagcacgttgtccgggcacaggtcaccgtggcaaaccgccaga
tcctcgtccgcaggccgagtccgctccagctcggcgagaagccgctcccccgaccacccttccgctcc
tcgtccagatcctccaagtcgacgctcccttcagcgacagcacgggccgcctgcggcaccgtcaccgcg
agactgcgatcgaacggacaccgctcccagtccagcgcgtgcagcgaacgagcgagccccgcgagcgcc
accgccacgtccagccgctgctcccgcggccaccgcgcactggccggacgccccggaaccgcttcggtg
accaaccaggcgaccctctcgtccccaccaccctccacaacacgaggtacgggaatccccacctccgcc
aaccacaccagccgctcagcctcacccaacaagcccacccccggcccccagagctgccaccttgacaaac
aactcccgcccaccaccccgaagccgataaacaccagcccccgaggccccatcctccacaacaacccac
tcacaaccgggataccgaccccgcagtgcacgcaacgcatcgtccatgcttcgaaattcttcagcaccg
gggagggcggagtggccatcctgcaaatggaaacggcgacgcagggttagatgctgcttgagacagcga
cagaggagccaaaagccttcgtcgacacaatgcgggcgttgcaagtcaaatctgcaagcacgctgcctg
atccgccgggcttgctcgtcgactcacctggccattttaagatgttgagtgacttctcttgtaaaaaag
taaagaacataggcccccctggccggtttatcaggagggcaccgctccaggggctgcatgcgaactgctt
gcattggcgcctagcctttgtgggccaggggcttccggataagggttgcaagtgctcaaataccccat
caaacatcatcctggtttggctgcgctccttctggcgcgcccggcatgcaagcttgatgggatcttaag
ctagcaattccagcctgccctgttcatgcctctggcattgaactttattctcatcgatgctgttgtagc
tttacagtgaccaagatggacaggtgtacagtgggcgttgcatgaaagggtggaagtgatgaggaaatc
aaaattgggtactaggtcagggaaggtggaggtcgtgggcaaaaggggtgctccggatgcatggtggg
tttgagaggagtaaagtataatctgacagccgatggggcaacaattgcgtgtgagacagtgaagtcaga
ggactccatcgcgcatgaagcaagagagccttccgaggccctggaacactcgaacagcatcctgcagga
tccttatatttctctgactaggcagcattggttatttgccgtgccagggcccactctagcccagactcc
ccagactccagatcgaggtcagcagcgtcctcggagtccctctcctccgcagcagtaaggcagtcaacc

gactgcagcttgcggagactgttgaggcctgcacatgtgttggttgccagagggagtgacaaggggtga
aaggtaaggatgtacaaggcacgcaccatggcaggcgattgcagccatatttcaagggcatcttgaaac
tgctgtcgtgggaaactctcactggccccgctctcccagctctactgtaaagttgccagatcccaccat
cagctcaccattctggtctgaggtcgccaccaccgacgtctgcgcggcattgaaggcgtctgcagccgc
ctgggcggcagcgacgttgcgcatggtgtcatccaaatcgctctgctcgatcacctccccctcatcatc
atacatgggaatgtcctcgtcctgtattatggggtgcatttatggcccgtagggcagacagcacgcaag
gaaaggacaaagacgcgacggcagcgagtcaccacaacctgcaccaatgtcaacagtgctaagattctg
gtcctcgctctccatccaccactccggggtcgcatatccgacgccccaggaatcgttcccaaatccact
ctccatcgagttcccggctgaggcagggcggtaccagatctgtataccatatgttaattaagagctctc
tagagtcgacctgcaggggggcagtgagttatcaggaacagggtggagcctgtaaggggatgcatctca
gcggctcaagccgaggacatctgaggccggaacctccacagtttctactctctccctgtatgatgagca
gctgcacgacccttttgtcgagaagaatatgtgtaccgaggatgtggtgctgtcaacaacacaacactgc
gaggaagggggccaggggcatgtttgtacaggggcgtgcaccgccaggcatctgcatggtgccgagtca
ccagcatgtcagaggcagctgctattggggaaagggcttgggatacccggggtttgttgtgcatgctta
tcgtgcatgtggcacagggcagcagcaagccgggtctggtcagggaaggcttcctggggtccttgagat
gggctttgccacatccagtcccctgcacgcctcgaccccgggccctggggcccgtccacctacccctcc
tccccaaactcggttgtgaacttcttgtacgtctccaggtcgtcgaattccagcctgccctgttcatgc
ctctggcattgaactttattctcatcgatgctgttgtagctttacagtgaccaagatggacaggtgtac
agtgggcgttgcatgaaagggtggaagtgatgaggaaatcaaaattgggtactaggtcagggaaggtgg
aggtcgtgggcaaaaggggtgctccggatgcatggtgggtttgagaggagtaaagtataatctgacag
ccgatggggcaacaattgcgtgtgagacagtgaagtcagaggactccatcgcgcatgaagcaagagagc
cttccgaggccctggaacactcgaacagcatcctgcaggatccttatatttctctgactaggcagcatt
ggttatttgccgtgccagggcccactctagcccagactccccagactccagatcgaggtcagcagcgtc
ctcggagtccctctcctccgcagcagtaaggcagtcaaccgactgcagcttgcggagactgttgaggcc
tgcacatgtgttggttgccagagggagtgacaaggggtgaaaggtaaggatgtacaaggcacgcaccat
ggcaggcgattgcagccatatttcaagggcatcttgaaactgctgtcgtgggaaactctcactggcccc
gctctcccagctctactgtaaagttgccagatcccaccatcagctcaccattctggtctgaggtcgcca
ccaccgacgtctgcgcggcattgaaggcgtctgcagccgcctgggcggcagcgacgttgcgcatggtgt
catccaaatcgctctgctcgatcacctccccctcatcatcatacatgggaatgtcctcgtcctgtatta
tggggtgcatttatggcccgtagggcagacagcacgcaaggaaaggacaaagacgcgacggcagcgagt
caccacaacctgcaccaatgtcaacagtgctaagattctggtcctcgctctccatccaccactccgggg
tcgcatatccgacgccccaggaatcgttcccaaatccactctccatcgagttcccggctgaggcagggc
gaatgctaatggacgatgcgttgcgtgcactgcggggtcggtatcccggttgtgagtgggttgttgtgg
aggatggggcctcgggggctggtgtttatcggcttcggggtggtgggcgggagttgtttgtcaaggtgg
cagctctgggggccggggtgggcttgttgggtgaggctgagcggctggtgtggttggcggaggtgggga
ttcccgtacctcgtgttgtggagggtggtggggacgagagggtcgcctggttggtcaccgaagcggttc
cggggcgtccggccagtgcgcggtggccgcgggagcagcggctggacgtggcggtggcgctcgcggggc
tcgctcgttcgctgcacgcgctggactgggagcggtgtccgttcgatcgcagtctcgcggtgacggtgc
cgcaggcggcccgtgctgtcgctgaagggagcgtcgacttggaggatctggacgaggagcggaaggggt
ggtcgggggagcggcttctcgccgagctggagcggactcggcctgcggacgaggatctggcggtttgcc
acggtcacctgtgcccggacaacgtgctgctcgaccctcgtacctgcgaggtgaccgggctgatcgacg
tggggcgggtcggccgtgcggaccggcactccgatctcgcgctggtgctgcgcgagctggcccacgagg
aggacccgtggttcgggccggagtgttccgcggcgttcctgcgggagtacgggcgcgggtgggatgggg
cggtatcggaggaaaagctggcgttttaccggctgttggacgagttcttctgac

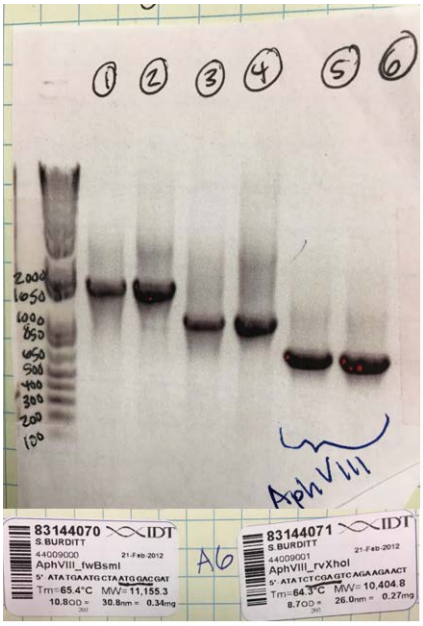


Figure 10: PCR gel and primers confirming insertion of AphVIII into pCpAphVIII plasmid.

Following an extensive genome exploration for optimal promoters and terminators for *C. sorokiniana* LANL researchers further engineered pCpAphVIII to create the pCSI plasmid shown in Figure 11 by synthesizing an insert (CSI promoters/terminators insert) containing two native *C. sorokiniana* promoter/terminator pairs: 1) the actin promoter/terminator pair, and 2) psaD promoter/terminator pair. The sequence of the CSI promoters/terminators insert is shown in Figure 12 and was cloned into pCpAphVIII at the SapI and NotI restriction sites (Figure 9) by digestion/ligation. The KpnI and NdeI restriction digest gel confirming the insertion of the CSI promoters/terminators insert in pCSI is shown in Figure 13 and the final pCSI plasmid sequence is shown in Figure 14.

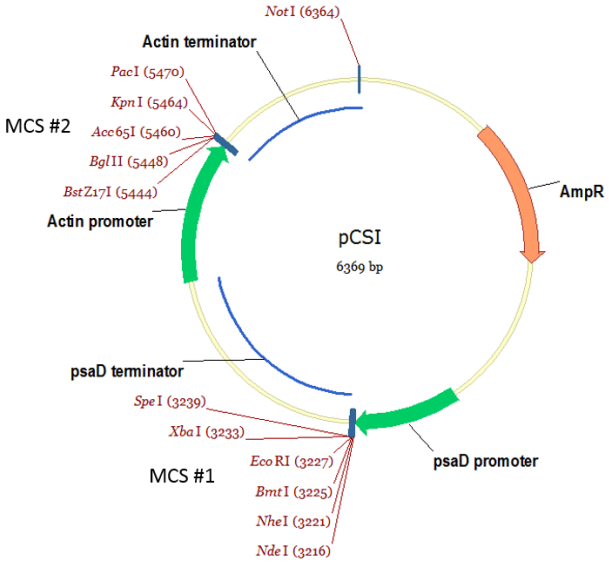


Figure 11: Map of pCSI plasmid showing key restriction enzyme sites. psaD = native *C. sorokiniana* photosynthesis-related gene promoter/terminator pair; bla = Ampicillin resistance gene; MCS = Multiple Cloning Site

Figure 12: Sequence of *C. sorokiniana* promoter/terminator pairs inserted into pCpAphVIII at the SapI and NotI restriction sites (Figure 9) to create pCSI (Figure 11). Green highlight = native *C. sorokiniana* psaD promoter sequence; Red highlight = native *C. sorokiniana* psaD terminator sequence; Grey Highlight = native *C. sorokiniana* Actin promoter sequence; Pink Highlight = native *C. sorokiniana* Actin terminator sequence; Yellow = Multiple Cloning Sites (MCS); Blue =SapI, NotI and ClaI restriction enzyme sites.

GCGCTCTTCCGCTTTTCAGCAGGTTATTGGTTCATGCACGGTCATGTTCCTGGTCGCTTGTATCTGCGA
TCTGCAACGATATCCATGATGACACGCAAGTCCACGCATTGAGCAGCTCCGTGGTGCCACAGCTGGT
CCAGCCTTGCTACAAGTGCGGGTAGAGCAACAGCACAGCCGTGGCGGCTGGAATTTGCCATGGCCT
GTCCATATCTACCAACCAGCGCGGGATCTGTGAAGGGCATGTGAGCGGCGCCGCACGGGCGGCGGC
TGCCGCCAACCGGCTGTCGCCGCCTGCTGCAGCGGGTCGCCTGCACCCCTTGCAGTGACTTGCTGTT
CACATCAAGAGCAATGAATGAATGCCCTGCCGCGGTGAATGCCCGCAGCTGCTGCCCCAGGCAGTC
GGGGGGACGGCACCCCAGCCGGCAGGTCGCCCGGCGCGGATTGCCGGCGGCTTTCTGAATGCAGG
CACTGCCGATAAAGTGGTTGGGCTCCCACACAGCACTTTGCAGCGCCACCCTGCTGAGCATTTCACC
GGGTCACAGCAGGATCGTCAGCGAGAAAGGCTCCCCCCCAGCGACCGCTCGGCCGCTGCTCCCACC
CCCATATAGCCCCAGCCGACCAGGACAGAGGCCAAACCATATGGCTAGCGAATTCTCTAGAACTAGT
CTCATTAAATGCTGCTGCTGTTTTTGTTGCAAACCGCTAGATTTGCACACCCCTCTCTCCCGGCCGTGT
AAGGCTGAACCCAACACGTGGCACCGGGGATTGTTGCGAGCCTGGCTGGGGAAGACATGGGCAGCT
GGTTGAGATGTTGAATTACAGCTGTACATCAGTTTGCGTGCATTGTGTCCGAACCTGGAGCACTGCAA
GGCACTGGCGGCTGGAAGTGGCTGGTCAGTTCACTCTTCATTCACCAGCACCACCTCCTCCTGCACC
AGGGCAGCCTCCCCTCCGCAGCTGCCGCCGCAGCTGAACACAGCCACGGTGAGCCACTCCCAGCTG
TCGGGCGGCTGCAGGGAGGCTGCCGCTGCTGCTCCGCCCTCCTCAGCTGCCAGCCATTCGGCTGCCT
CGGTCAAAGCGGGGATGAGGGGTGTCATCAGCTGCAGCTCAAACACGCGCGGCCGGCCGCAGGCG
GCACAGGGTGCGGGAGCGGGCTGCTCGCGAGTCGGCCACAGCGGTGCACCGCCAAAGCTGCAGGG
TGAAGGGGCCAGAGGACAGCAGAGAAGTCGGTTTTGGGTCACTTCAGTAGTCACCAGCGGGCAGCA
AAACTGACCTCCAGCCCTCTTGCTTGCACTGACCACCCCTTGGCGGCCCCGCTCGCCTGTAGCGCGC
GCACTGGTCGGGGCAGCGGCCCAGCTGCTTTGCAAACTTGAAGAAGGCGGCTCCCACGCCCGCCCG
CGTGCCGCCGGTGGGCGCCAGCACAGCGTCCTCTTCATAGCCTTCTCCTGACCAGGACTCAGGGCCG
CCCTCCACGCTGGGCGCACATGCTGTGCCCGCTGTGGTCGCCATGGCGGCGCCCGCCTCGCCAGCT
GCTGCCGCCTCTGCCTCGTATCGAGCCATCAGCTGCTGCAGGTGCTCCTGTTCGGAGCGCTGCTGCC
TGCCGCTGCTGGCGCTGCACTGCGACTCCGGCTCGGCATACAGGTAGAAGGCGGGGAGGCTGGGCT
GGGTTGGGTCATACGCTGGCGCTGCCCCAGCTGCTGCCACCTCGCTGCTGCGCAGCTCCCCCTGCTG
CTGCGGCTGCTGCTTGCCGGCTGGCTTGGCAGCTGGGGCAGCACTGCCCACCGCTTCCAGCGCTGC
GTTGAGGTCTCCAAAGTCAAACGGAGCTGCCGCACTGTCCGTCACATCAGCGCCCATCCCCCAGTCA
TCGCTGCCTCCTGCACCCCAGTCATCTGCACCGCCCATACCCCAGTCATCGGCTGCAGGCGGCTTCT
CTGCCGGCGGCGCTGCCTGCTGCTGCAGCTGCTGCGCTGTCGGCGGCTGCTCGCTGCTGGCTGGAGT
GCTGGGCAGGGTATCGATGGCGGCAGCCAGCTCAAGTGTGTGTGGCTCTGGGCTGCTGCCGGTGTTG
ACCGCCGCTGCGCCGCCGCCCGGCAGCCAGTGCACCAGCAGCTGCGGGCCGGCGGACAGGCACTT
GACATACAGCGGCGAACGGCTGCCTGCAAGCGGGAATCGGTTGAGGTCAGCTACGGTGCCTCGTGA
CGGCAGCCAGCAGCGATCAGCAAGCCACGGCTGGGAGCCATGGCAGCCACCCGACCAACACCCAA
CAACCCAGCGGCAACCCAGGTCAGCAGCCGCTGCACGCGCCTTGCGAGTCTTGGTAGCGAAAGGCG
TAGGAACCCTCCAGAGATTCCCACCCAGAGGTTCCCACTTCCTCACGGTCGGCGCCCCAGTCTGCAC
ATGTGGATGGGGGGTGGGGCAATTAGTAGCGAGGATCGGTTAGGCAGGGCTGGCACGCGAGGGCAA
AGGGCCGAACAGGGCTCTTGGAAAGCTGATCATGCACCCAGCAAGTGCAATCGGCCACTGTGATTGG
CTGCTTCTCTAAACATCTGATCCTGATCCAGCAGACAAAGCCAAGATCCGGCCTCTCACCGGCACCCT
CGTCCTCGGCGGCAGGGCCGGCGGCCACCAGCTTGTAGCCTTCGGCCAGCAGGAAGGCGTGCACCG
CAAACGCCAGCTTGTCGGCGTCGGACCTGGCGAGCATGAGAAGATACAACTTGGTTGACGATGTGCA
GGTGTGCTGGACTGGCCGAGGGGCTGGACGGCCGAGGGGCCCGCAGGCGGCTGGCGGCGATGGCG
GCGGCGCGGCGGCGGTGCCAGCAGCAGCGCTTTGCGGCCCGGCACCTTGCTGACCTGAAGTTGGGG
CGCGGTATACAGATCTAAGCTTGGTACCTTAATTAAGAGCTCCTCGGGCCGGCAGTAGTTGCTACCTG
GGAGTGGTCGGCGCGGTGTGCACTTGAAGCAGGGGCGGCTTGCGTAGTGCAGCAGCGTGGGAGGAA
ATTGCAAAGACTGTGCAGCCCACCTCATCTAGCGCCTGGGATCGCATGCACACTTCACATGCTGGCA
ATGGTCTAGGCCCTGCCCTTTTCCTGGAGGCTGCTCTGCCATACACCTTGCCTGACATTACAGCATGG
ACTGTTGCATCGGGTCGACAGCGCACATGCAAGCCGAGCCACACCGCTGCATTCGAGCATCAGGCC
AGCACCATGGGCTGCCCAGGCAGGCAGCCTAGCGTCCGCCCAGCTTCTGCCGCAGGGCTTGCAGGG
CATCCGGCCCGCTGGCCTCCCGGCCCAGCGCAGTGGCAACCACAGTCATGCCACGGCCGCCCGCTT
TCTTGCGCTGCTTGGCTGTCAGGGCCTTGTTGAGTGCCGCCTGCAGGGCCGCCGCTCGAGGGTCAAC

AACCTTCTTTTTGGCGGCCTTCTTCTTCTGCTGCTTGGCGCCCGATGAGCCCGCTGTGCCAGCTGCGC
CCGCTGCACTCGCCCGGCTGCGGCTGTCCGACTGCCGCTGAGCCAGAGCCACAGCCTGGCTCACAG
TCTGCACTGTGCGAGGCTCCGGCACAGGTGCCGCTGCCTCGCCTGCCGCTGCCGCACCCTCTGCTGC
TGCCGCCGCTGCTTTCTTGCGCTTCTTGCCACCGGCTGCTGCGGGGGCCGAAGGGTCGACTGCCGCC
ATGACTGCTGCTGCTGCGCCGCCGGTGAACCACTTGCGGCTGAACTTGGGCTGCAGCGGCTCGGCCA
GCAGCGCTGCCAGCTGTGCACGCACATCTGCTGCCGTCTCTGCAGCACAGCATTGTGCGGCCGC

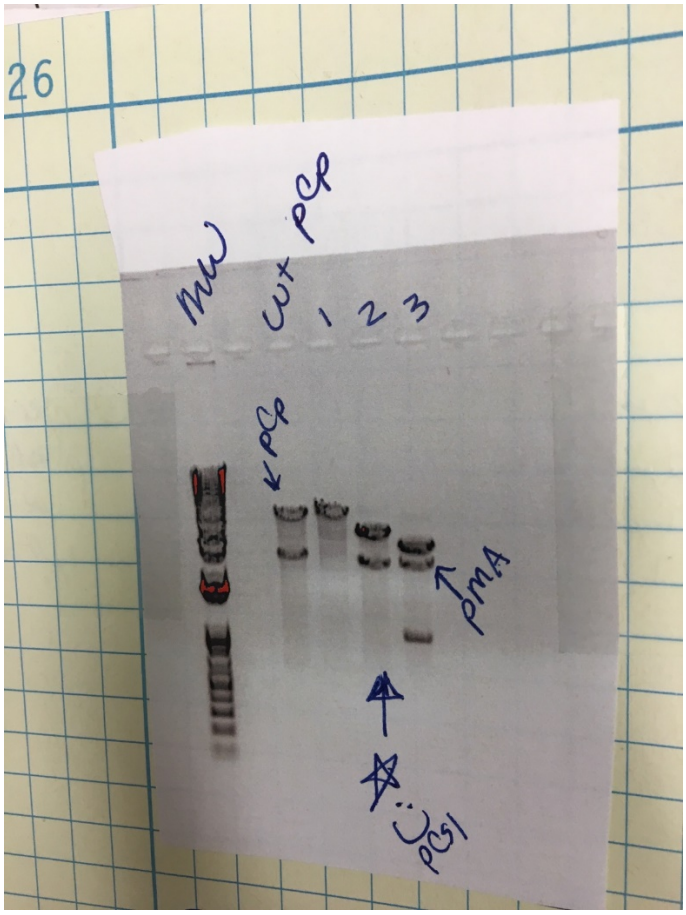


Figure 13: KpnI (pCSI cut site 5464 bp) and NdeI (pCSI cut site 3216 bp) restriction digest gel confirming insertion of the CSI promoters/terminators insert into the pCpAphVIII backbone at SapI/NotI (Figure 9) to create pCSI (Figure 11).

Figure 14: Sequence of pCSI plasmid (map shown in Figure 11). Lowercase lettering = pCPAphVIII backbone; Uppercase lettering = pCSI insert (Figure 12).

caccgcggtggagctccaattcgccctatagtgagtcgtattacaattcactggccgtcgttttacaac
gtcgtgactgggaaaaccctggcgttacccaacttaatcgccttgcagcacatccccctttcgccagct
ggcgtaatagcgaagaggcccgcaccgatcgcccttcccaacagttgcgcagcctgaatggcgaatgga
aattgtaagcgttaatattttgttaaaattcgcgttaaattttgttaaatcagctcatttttaacca
ataggccgaaatcggcaaaatcccttataaatcaaaagaatagaccgagatagggttgagtgttgttcc
agtttggaacaagagtccactattaaagaacgtggactccaacgtcaaagggcgaaaaaccgtctatca
gggcgatggcccactacgtgaaccatcaccctaatcaagtttttggggtcgaggtgccgtaaagcact
aaatcggaaccctaaagggagcccccgatttagagcttgacggggaaagccggcgaacgtggcgagaaa
ggaagggaagaaagcgaaaggagcgggcgctagggcgctggcaagtgtagcggtcacgctgcgcgtaac
caccacacccgccgcgcttaatgcgccgctacagggcgcgtcaggtggcactttcggggaaatgtgcg
cggaacccctatttgtttatttttctaaatacattcaaatatgtatccgctcatgagacaataaccctg

ataaatgcttcaataatattgaaaaaggaagagtatgagtattcaacatttccgtgtcgcccttattcc
ctttttgcggcattttgccttcctgtttttgctcacccagaaacgctggtgaaagtaaaagatgctga
agatcagttgggtgcacgagtgggttacatcgaactggatctcaacagcggtaagatccttgagagttt
tcgccccgaagaacgttttccaatgatgagcacttttaaagttctgctatgtggcgcggtattatcccg
tattgacgccgggcaagagcaactcggtcgccgcatacactattctcagaatgacttggttgagtactc
accagtcacagaaaagcatcttacggatggcatgacagtaagagaattatgcagtgctgccataaccat
gagtgataacactgcggccaacttacttctgacaacgatcggaggaccgaaggagctaaccgcttttt
gcacaacatgggggatcatgtaactcgccttgatcgttgggaaccggagctgaatgaagccataccaaa
cgacgagcgtgacaccacgatgcctgtagcaatggcaacaacgttgcgcaaactattaactggcgaact
acttactctagcttcccggcaacaattaatagactggatggaggcggataaagttgcaggaccacttct
gcgctcggcccttccggctggctggtttattgctgataaatctggagccggtgagcgtgggtctcgcgg
tatcattgcagcactggggccagatggtaagccctcccgtatcgtagttatctacacgacggggagtca
ggcaactatggatgaacgaaatagacagatcgctgagataggtgcctcactgattaagcattggtaact
gtcagaccaagtttactcatatatactttagattgatttaaaacttcatttttaatttaaaaggatcta
ggtgaagatccttttgataatctcatgaccaaaatcccttaacgtgagttttcgttccactgagcgtc
agaccccgtagaaaagatcaaaggatcttcttgagatccttttttttctgcgcgtaatctgctgcttgca
aacaaaaaaaccaccgctaccagcggtggtttgtttgccggatcaagagctaccaactctttttccgaa
ggtaactggcttcagcagagcgcagataccaaatactgttcttctagtgtagccgtagttaggccacca
cttcaagaactctgtagcaccgcctacatacctcgctctgctaatcctgttaccagtggctgctgccag
tggcgataagtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcgcagcggtcggg
ctgaacgggggttcgtgcacacagcccagcttggagcgaacgacctacaccgaactgagatacctaca
gcgtgagctatgagaaagcgccacgcttcccgaagggagaaaggcggacaggtatccggtaagcggcag
ggtcggaacaggagagcgcacgagggagcttccagggggaaacgcctggtatctttatagtcctgtcgg
gtttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcaggggggcggagcctatggaaaaa
cgccagcaacgcggccttttacggttcctggccttttgctggccttttgctcacatgttctttcctgc
gttatcccctgattctgtggataaccgtattaccgcctttgagtgagctgataccgctcgccgcagccg
aacgaccgagcgcagcgagtcagtgagcgaggaaGCTTTTCAGCAGGTTATTGGTTCATGCACGGTCAT
GTTCCTGGTCGCTTGTATCTGCGATCTGCAACGATATCCATGATGACACGCAAGTCCACGCATTGAGCA
GCTCCGTGGTGCCACAGCTGGTCCAGCCTTGCTACAAGTGCGGGTAGAGCAACAGCACAGCCGTGGCGG
CTGGAATTTGCCATGGCCTGTCCATATCTACCAACCAGCGCGGGATCTGTGAAGGGCATGTGAGCGGCG
CCGCACGGGCGGCGGCTGCCGCCAACCGGCTGTCGCCGCCTGCTGCAGCGGGTCGCCTGCACCCCTTGC
AGTGACTTGCTGTTCACATCAAGAGCAATGAATGAATGCCCTGCCGCGGTGAATGCCCGCAGCTGCTGC
CCCAGGCAGTCGGGGGGACGGCACCCCAGCCGGCAGGTCGCCCGGCGCGGATTGCCGGCGGCTTTCTGA
ATGCAGGCACTGCCGATAAAGTGGTTGGGCTCCCACACAGCACTTTGCAGCGCCACCCTGCTGAGCATT
TCACCGGGTCACAGCAGGATCGTCAGCGAGAAAGGCTCCCCCCCAGCGACCGCTCGGCCGCTGCTCCCA
CCCCCATATAGCCCCAGCCGACCAGGACAGAGGCCAAACCATATGGCTAGCGAATTCTCTAGAACTAGT
CTCATTAAATGCTGCTGCTGTTTTTGTTGCAAACCGCTAGATTTGCACACCCCTCTCTCCCGGCCGTGT
AAGGCTGAACCCAACACGTGGCACCGGGGATTGTTGCGAGCCTGGCTGGGGAAGACATGGGCAGCTGGT
TGAGATGTTGAATTACAGCTGTACATCAGTTTGCGTGCATTGTGTCCGAACCTGGAGCACTGCAAGGCA
CTGGCGGCTGGAAGTGGCTGGTCAGTTCACTCTTCATTCACCAGCACCACCTCCTCCTGCACCAGGGCA
GCCTCCCCTCCGCAGCTGCCGCCGCAGCTGAACACAGCCACGGTGAGCCACTCCCAGCTGTCGGGCGGC
TGCAGGGAGGCTGCCGCTGCTGCTCCGCCCTCCTCAGCTGCCAGCCATTCGGCTGCCTCGGTCAAAGCG
GGGATGAGGGGTGTCATCAGCTGCAGCTCAAACACGCGCGGCCGGCCGCAGGCGGCACAGGGTGCGGGA
GCGGGCTGCTCGCGAGTCGGCCACAGCGGTGCACCGCCAAAGCTGCAGGGTGAAGGGGCCAGAGGACAG
CAGAGAAGTCGGTTTTGGGTCACTTCAGTAGTCACCAGCGGGCAGCAAAACTGACCTCCAGCCCTCTTG
CTTGCACTGACCACCCCTTGGCGGCCCCGCTCGCCTGTAGCGCGCGCACTGGTCGGGGCAGCGGCCCAG
CTGCTTTGCAAACTTGAAGAAGGCGGCTCCCACGCCCGCCCGCGTGCCGCCGGTGGGCGCCAGCACAGC
GTCCTCTTCATAGCCTTCTCCTGACCAGGACTCAGGGCCGCCCTCCACGCTGGGCGCACATGCTGTGCC
CGCTGTGGTCGCCATGGCGGCGCCCGCCTCGCCAGCTGCTGCCGCCTCTGCCTCGTATCGAGCCATCAG
CTGCTGCAGGTGCTCCTGTTCGGAGCGCTGCTGCCTGCCGCTGCTGGCGCTGCACTGCGACTCCGGCTC
GGCATACAGGTAGAAGGCGGGGAGGCTGGGCTGGGTTGGGTCATACGCTGGCGCTGCCCCAGCTGCTGC

CACCTCGCTGCTGCGCAGCTCCCCCTGCTGCTGCGGCTGCTGCTTGCCGGCTGGCTTGGCAGCTGGGGC
AGCACTGCCCACCGCTTCCAGCGCTGCGTTGAGGTCTCCAAAGTCAAACGGAGCTGCCGCACTGTCCGT
CACATCAGCGCCCATCCCCCAGTCATCGCTGCCTCCTGCACCCCAGTCATCTGCACCGCCCATACCCCA
GTCATCGGCTGCAGGCGGCTTCTCTGCCGGCGGCGCTGCCTGCTGCTGCAGCTGCTGCGCTGTCGGCGG
CTGCTCGCTGCTGGCTGGAGTGCTGGGCAGGGTATCGATGGCGGCAGCCAGCTCAAGTGTGTGTGGCTC
TGGGCTGCTGCCGGTGTTGACCGCCGCTGCGCCGCCGCCCGGCAGCCAGTGCACCAGCAGCTGCGGGCC
GGCGGACAGGCACTTGACATACAGCGGCGAACGGCTGCCTGCAAGCGGGAATCGGTTGAGGTCAGCTAC
GGTGCCTCGTGACGGCAGCCAGCAGCGATCAGCAAGCCACGGCTGGGAGCCATGGCAGCCACCCGACCA
ACACCCAACAACCCAGCGGCAACCCAGGTCAGCAGCCGCTGCACGCGCCTTGCGAGTCTTGGTAGCGAA
AGGCGTAGGAACCCTCCAGAGATTCCCACCCAGAGGTTCCCACTTCCTCACGGTCGGCGCCCCAGTCTG
CACATGTGGATGGGGGGTGGGGCAATTAGTAGCGAGGATCGGTTAGGCAGGGCTGGCACGCGAGGGCAA
AGGGCCGAACAGGGCTCTTGGAAAGCTGATCATGCACCCAGCAAGTGCAATCGGCCACTGTGATTGGCT
GCTTCTCTAAACATCTGATCCTGATCCAGCAGACAAAGCCAAGATCCGGCCTCTCACCGGCACCCTCGT
CCTCGGCGGCAGGGCCGGCGGCCACCAGCTTGTAGCCTTCGGCCAGCAGGAAGGCGTGCACCGCAAACG
CCAGCTTGTCGGCGTCGGACCTGGCGAGCATGAGAAGATACAACTTGGTTGACGATGTGCAGGTGTGCT
GGACTGGCCGAGGGGCTGGACGGCCGAGGGGCCCGCAGGCGGCTGGCGGCGATGGCGGCGGCGCGGCGG
CGGTGCCAGCAGCAGCGCTTTGCGGCCCGGCACCTTGCTGACCTGAAGTTGGGGCGCGGTATACAGATC
TAAGCTTGGTACCTTAATTAAGAGCTCCTCGGGCCGGCAGTAGTTGCTACCTGGGAGTGGTCGGCGCGG
TGTGCACTTGAAGCAGGGGCGGCTTGCGTAGTGCAGCAGCGTGGGAGGAAATTGCAAAGACTGTGCAGC
CCACCTCATCTAGCGCCTGGGATCGCATGCACACTTCACATGCTGGCAATGGTCTAGGCCCTGCCCTTT
TCCTGGAGGCTGCTCTGCCATACACCTTGCCTGACATTACAGCATGGACTGTTGCATCGGGTCGACAGC
GCACATGCAAGCCGAGCCACACCGCTGCATTCGAGCATCAGGCCAGCACCATGGGCTGCCCAGGCAGGC
AGCCTAGCGTCCGCCCAGCTTCTGCCGCAGGGCTTGCAGGGCATCCGGCCCGCTGGCCTCCCGGCCCAG
CGCAGTGGCAACCACAGTCATGCCACGGCCGCCCGCTTTCTTGCGCTGCTTGGCTGTCAGGGCCTTGTT
GAGTGCCGCCTGCAGGGCCGCCGCTCGAGGGTCAACAACCTTCTTTTTGGCGGCCTTCTTCTTCTGCTG
CTTGGCGCCCGATGAGCCCGCTGTGCCAGCTGCGCCCGCTGCACTCGCCCGGCTGCGGCTGTCCGACTG
CCGCTGAGCCAGAGCCACAGCCTGGCTCACAGTCTGCACTGTGCGAGGCTCCGGCACAGGTGCCGCTGC
CTCGCCTGCCGCTGCCGCACCCTCTGCTGCTGCCGCCGCTGCTTTCTTGCGCTTCTTGCCACCGGCTGC
TGCGGGGGCCGAAGGGTCGACTGCCGCCATGACTGCTGCTGCTGCGCCGCCGGTGAACCACTTGCGGCT
GAACTTGGGCTGCAGCGGCTCGGCCAGCAGCGCTGCCAGCTGTGCACGCACATCTGCTGCCGTCTCTGC
AGCACAGCATTGTGCGGCCGC

Next, the AphVIII gene was codon-optimized for C. sorokiniana and synthesized for insertion into the KpnI and BglII restriction enzyme sites in the 2nd Multiple Cloning Site (MCS) of the pCSI vector (Figure 11). The cloned AphVII sequence is shown in Figure 15 and the gel of the restriction digestions used to confirm insertion of AphVIII into pCSI is shown in Figure 16. The resulting vector, pCSIA2 (Figure 17; sequence in Figure 17A), was sequence confirmed by Genewiz using the primers in Figure 18.

In an attempt to improve gene expression in strains transformed with the pCSIA2 vector, the psaD terminator sequence was extended and the MCS between psaD promoter and terminator was modified resulting in an addition of 354 nucleotides. The construct for making these modification (Figure 19) was synthesized and cloned into pCSIA2 by Genewiz to create the PACE Chlorella plasmid shown in Figure 2 of this TERA application.

Figure 15: Sequence of AphVIII codon-optimized for C. sorokiniana for insertion into pCSI to create the pCSIA2 plasmid (Figure 16 and 17)

AGATCTCATATGATGGACGACGCCCTGCGCGCCCTGCGCGGCCGCTACCCCGGCTGCGAGTGGGTGGTG
GTGGAGGACGGCGCCAGCGGCGCCGGCGTGTACCGCCTGCGCGGCGGCGGCCGCGAGCTGTTCGTGAAG
GTGGCCGCCCTGGGCGCCGGCGTGGGCCTGCTGGGCGAGGCCGAGCGCCTGGTGTGGCTGGCCGAGGTG
GGCATCCCCGTGCCCCGCGTGGTGGAGGGCGGCGGCGACGAGCGCGTGGCCTGGCTGGTGACCGAGGCC
GTGCCCGGCCGCCCCGCCAGCGCCCGCTGGCCCCGCGAGCAGCGCCTGGACGTGGCCGTGGCCCTGGCC

GGCCTGGCCCGCAGCCTGCACGCCCTGGACTGGGAGCGCTGCCCCTTCGACCGCAGCCTGGCCGTGACC
GTGCCCCAGGCCGCCCGCGCCGTGGCCGAGGGCAGCGTGGACCTGGAGGACCTGGACGAGGAGCGCAAG
GGCTGGAGCGGCGAGCGCCTGCTGGCCGAGCTGGAGCGCACCCGCCCCGCCGACGAGGACCTGGCCGTG
TGCCACGGCCACCTGTGCCCCGACAACGTGCTGCTGGACCCCCGCACCTGCGAGGTGACCGGCCTGATC
GACGTGGGCCGCGTGGGCCGCGCCGACCGCCACAGCGACCTGGCCCTGGTGCTGCGCGAGCTGGCCCAC
GAGGAGGACCCCTGGTTCGGCCCCGAGTGCAGCGCCGCCTTCCTGCGCGAGTACGGCCGCGGCTGGGAC
GGCGCCGTGAGCGAGGAGAAGCTGGCCTTCTACCGCCTGCTGGACGAGTTCTTCTGAGAATTCGGTACC

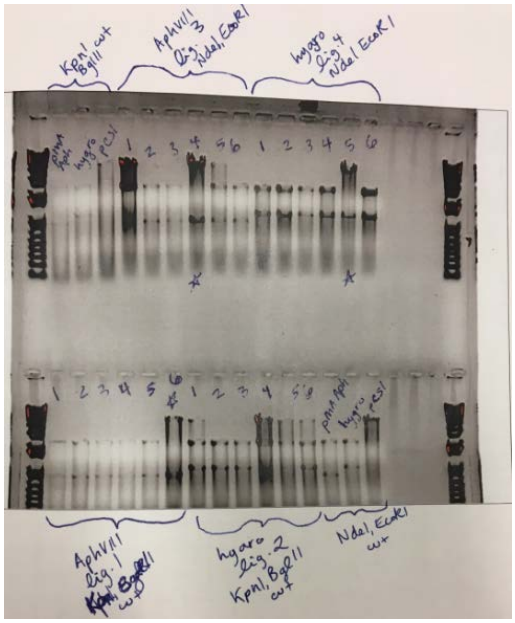


Figure 16: Gel of restriction digests to confirm insertion of AphVIII at MCS2 of pCSI to create pCSIA2. Constructs 1-6, among others, had the correct banding pattern and were sequence confirmed to be pCSIA2 by Genewiz using the primers listed in Figure 18.

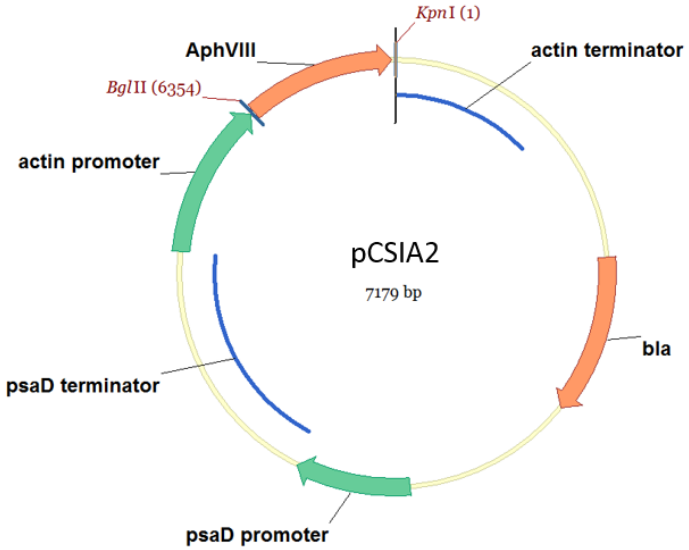


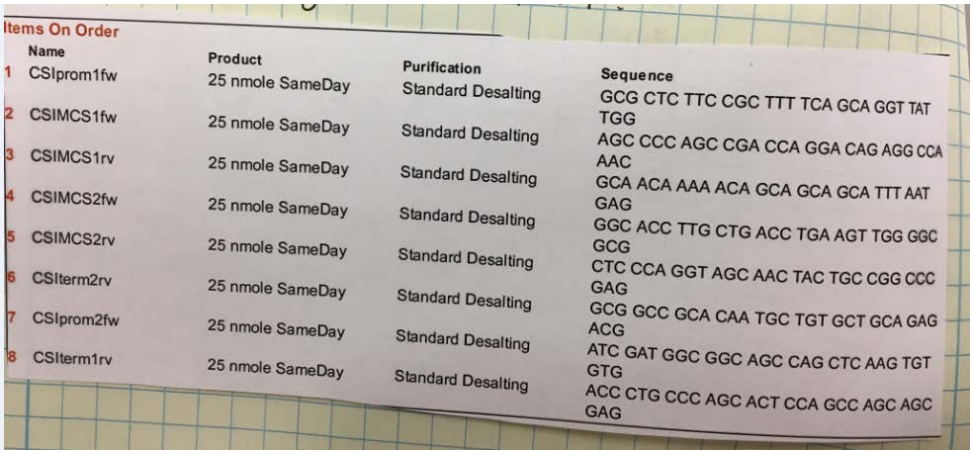
Figure 17: Map of pCSIA2 plasmid showing key restriction enzyme sites. psaD = native *C. sorokiniana* photosynthesis-related gene promoter/terminator pair; bla = Ampicillin resistance gene; AphVIII = Paromycin resistance gene.

Figure 17A: Sequence of pCSIA2 plasmid (map shown in Figure 17).

cttaattaagagctcctcgggccggcagtagttgctacctgggagtggtcggcgcggtgtgcacttgaa
gcaggggcggcttgcgtagtgcagcagcgtgggaggaaattgcaaagactgtgcagcccacctcatcta
gcgcctgggatcgcatgcacacttcacatgctggcaatggtctaggccctgccctttcctggaggctg
ctctgccatacaccttgcctgacattacagcatggactgttgcatcgggtcgacagcgcacatgcaagc
cgagccacaccgctgcattcgagcatcaggccagcaccatgggctgcccaggcaggcagcctagcgtcc
gcccagcttctgccgcagggcttgcagggcatccggcccgctggcctcccggcccagcgcagtggcaac
cacagtcatgccacggccgcccgctttcttgcgctgcttggctgtcagggccttgttgagtgccgcctg
cagggccgccgctcgagggtcaacaaccttctttttggcggccttcttcttctgctgcttggcgcccga
tgagcccgctgtgccagctgcgcccgctgcactcgcccggctgcggctgtccgactgccgctgagccag
agccacagcctggctcacagtctgcactgtgcgaggctccggcacaggtgccgctgcctcgcctgccgc
tgccgcaccctctgctgctgccgccgctgctttcttgcgcttcttgccaccggctgctgcgggggccga
agggtcgactgccgccatgactgctgctgctgcgccgccggtgaaccacttgcggctgaacttgggctg
cagcggctcggccagcagcgctgccagctgtgcacgcacatctgctgccgtctctgcagcacagcattg
tgcggccgccaccgcggtggagctccaattcgccctatagtgagtcgtattacaattcactggccgtcg
ttttacaacgtcgtgactgggaaaaccctggcgttacccaacttaatcgccttgcagcacatccccctt
tcgccagctggcgtaatagcgaagaggcccgcaccgatcgcccttcccaacagttgcgcagcctgaatg
gcgaatggaaattgtaagcgttaatattttgttaaaattcgcgttaaatttttgttaaatcagctcatt
ttttaaccaataggccgaaatcggcaaaatcccttataaatcaaaagaatagaccgagatagggttgag
tgttgttccagtttggaacaagagtccactattaaagaacgtggactccaacgtcaaagggcgaaaaac
cgtctatcagggcgatggcccactacgtgaaccatcaccctaatcaagtttttggggtcgaggtgccg
taaagcactaaatcggaaccctaaagggagcccccgatttagagcttgacggggaaagccggcgaacgt
ggcgagaaaggaagggaagaaagcgaaaggagcgggcgctagggcgctggcaagtgtagcggtcacgct
gcgcgtaaccaccacacccgccgcgcttaatgcgccgctacagggcgcgtcaggtggcacttttcgggg
aaatgtgcgcggaacccctatttgtttatttttctaaatacattcaaatatgtatccgctcatgagaca
ataaccctgataaatgcttcaataatattgaaaaaggaagagtatgagtattcaacatttccgtgtcgc
ccttattcccttttttgcggcattttgccttcctgtttttgctcacccagaaacgctggtgaaagtaaa
agatgctgaagatcagttgggtgcacgagtgggttacatcgaactggatctcaacagcggtaagatcct
tgagagttttcgccccgaagaacgttttccaatgatgagcacttttaaagttctgctatgtggcgcggt
attatcccgtattgacgccgggcaagagcaactcggtcgccgcatacactattctcagaatgacttggt
tgagtactcaccagtcacagaaaagcatcttacggatggcatgacagtaagagaattatgcagtgctgc
cataaccatgagtgataacactgcggccaacttacttctgacaacgatcggaggaccgaaggagctaac
cgcttttttgcacaacatgggggatcatgtaactcgccttgatcgttgggaaccggagctgaatgaagc
cataccaaacgacgagcgtgacaccacgatgcctgtagcaatggcaacaacgttgcgcaaactattaac
tggcgaactacttactctagcttcccggcaacaattaatagactggatggaggcggataaagttgcagg
accacttctgcgctcggcccttccggctggctggtttattgctgataaatctggagccggtgagcgtgg
gtctcgcggtatcattgcagcactggggccagatggtaagccctcccgtatcgtagttatctacacgac
ggggagtcaggcaactatggatgaacgaaatagacagatcgctgagataggtgcctcactgattaagca
ttggtaactgtcagaccaagtttactcatatatactttagattgatttaaaacttcatttttaatttaa
aaggatctaggtgaagatccttttgataatctcatgaccaaaatcccttaacgtgagttttcgttcca
ctgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatccttttttctgcgcgtaatctg
ctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgccggatcaagagctaccaactct
ttttccgaaggtaactggcttcagcagagcgcagataccaaatactgttcttctagtgtagccgtagtt
aggccaccacttcaagaactctgtagcaccgcctacatacctcgctctgctaatcctgttaccagtggc
tgctgccagtggcgataagtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcgca
gcggtcgggctgaacgggggttcgtgcacacagcccagcttggagcgaacgacctacaccgaactgag
atacctacagcgtgagctatgagaaagcgccacgcttcccgaagggagaaaggcggacaggtatccggt
aagcggcagggtcggaacaggagagcgcacgagggagcttccagggggaaacgcctggtatctttatag
tcctgtcgggtttcgccacctctgacttgagcgtcgattttgtgatgctcgtcaggggggcggagcct
atggaaaaacgccagcaacgcggccttttacggttcctggccttttgctggccttttgctcacatgtt

ctttcctgcgttatcccctgattctgtggataaccgtattaccgcctttgagtgagctgataccgctcg
ccgcagccgaacgaccgagcgcagcgagtcagtgagcgaggaagcttttcagcaggttattggttcatg
cacggtcatgttcctggtcgcttgtatctgcgatctgcaacgatatccatgatgacacgcaagtccacg
cattgagcagctccgtggtgccacagctggtccagccttgctacaagtgcgggtagagcaacagcacag
ccgtggcggctggaatttgccatggcctgtccatatctaccaaccagcgcgggatctgtgaagggcatg
tgagcggcgccgcacgggcggcggctgccgccaaccggctgtcgccgcctgctgcagcgggtcgcctgc
acccccttgcagtgacttgctgttcacatcaagagcaatgaatgaatgccctgccgcggtgaatgcccgc
agctgctgccccaggcagtcgggggacggcaccccagccggcaggtcgcccggcgcggattgccggcg
gctttctgaatgcaggcactgccgataaagtggttgggctcccacacagcactttgcagcgccaccctg
ctgagcatttcaccgggtcacagcaggatcgtcagcgagaaaggctccccccagcgaccgctcggccg
ctgctcccacccccatatagccccagccgaccaggacagaggccaaaccatatggctagcgaattctct
agaactagtctcattaaatgctgctgctgtttttgttgcaaaccgctagatttgcacacccctctctcc
cggccgtgtaaggctgaacccaacacgtggcaccggggattgttgcgagcctggctggggaagacatgg
gcagctggttgagatgttgaattacagctgtacatcagtttgcgtgcattgtgtccgaacctggagcac
tgcaaggcactggcggctggaagtggctggtcagttcactcttcattcaccagcaccacctcctcctgc
accagggcagcctcccctccgcagctgccgccgcagctgaacacagccacggtgagccactcccagctg
tcgggcggctgcagggaggctgccgctgctgctccgccctcctcagctgccagccattcggctgcctcg
gtcaaagcggggatgaggggtgtcatcagctgcagctcaaacacgcgcggccggccgcaggcggcacag
ggtgcgggagcgggctgctcgcgagtcggccacagcggtgcaccgccaaagctgcagggtgaaggggcc
agaggacagcagagaagtcggttttgggtcacttcagtagtcaccagcgggcagcaaaactgacctcca
gccctcttgcttgcactgaccacccctggcggccccgctcgcctgtagcgcgcgcactggtcggggca
gcggcccagctgctttgcaaacttgaagaaggcggctcccacgcccgcccgcgtgccgccggtgggcgc
cagcacagcgtcctcttcatagccttctcctgaccaggactcagggccgccctccacgctgggcgcaca
tgctgtgcccgctgtggtcgccatggcggcgcccgcctcgccagctgctgccgcctctgcctcgtatcg
agccatcagctgctgcaggtgctcctgttcggagcgctgctgcctgccgctgctggcgctgcactgcga
ctccggctcggcatacaggtagaaggcggggaggctgggctgggttgggtcatacgctggcgctgcccc
agctgctgccacctcgctgctgcgcagctccccctgctgctgcggctgctgcttgccggctggcttggc
agctggggcagcactgcccaccgcttccagcgctgcgttgaggtctccaaagtcaaacggagctgccgc
actgtccgtcacatcagcgcccatccccagtcatcgctgcctcctgcaccccagtcatctgcaccgcc
cataccccagtcatcggctgcaggcggcttctctgccggcggcgctgcctgctgctgcagctgctgcgc
tgtcggcggctgctcgctgctggctggagtgctgggcagggtatcgatggcggcagccagctcaagtgt
gtgtggctctgggctgctgccggtgttgaccgccgctgcgccgccgcccggcagccagtgcaccagcag
ctgcgggccggcggacaggcacttgacatacagcggcgaacggctgcctgcaagcgggaatcggttgag
gtcagctacggtgcctcgtgacggcagccagcagcgatcagcaagccacggctgggagccatggcagcc
acccgaccaacacccaacaacccagcggcaacccaggtcagcagccgctgcacgcgccttgcgagtctt
ggtagcgaaaggcgtaggaaccctccagagattcccacccagaggttcccacttcctcacggtcggcgc
cccagtctgcacatgtggatgggggtggggcaattagtagcgaggatcggttaggcagggctggcacg
cgagggcaaagggccgaacagggctcttggaaagctgatcatgcacccagcaagtgcaatcggccactg
tgattggctgcttctctaaacatctgatcctgatccagcagacaaagccaagatccggcctctcaccgg
caccctcgtcctcggcggcagggccggcggccaccagcttgtagccttcggccagcaggaaggcgtgca
ccgcaaacgccagcttgtcggcgtcggacctggcgagcatgagaagatacaacttggttgacgatgtgc
aggtgtgctggactggccgaggggctggacggccgaggggcccgcaggcggctggcggcgatggcggcg
gcgcggcggcggtgccagcagcagcgctttgcggcccggcaccttgctgacctgaagttggggcgcggt
atacagatctcatatgatggacgacgccctgcgcgccctgcgcggccgctaccccggctgcgagtgggt
ggtggtggaggacggcgccagcggcgccggcgtgtaccgcctgcgcggcggcggccgcgagctgttcgt
gaaggtggccgccctgggcgccggcgtgggcctgctgggcgaggccgagcgcctggtgtggctggccga
ggtgggcatccccgtgccccgcgtggtggagggcggcggcgacgagcgcgtggcctggctggtgaccga
ggccgtgcccggccgccccgccagcgcccgctggccccgcgagcagcgcctggacgtggccgtggccct
ggccggcctggcccgcagcctgcacgccctggactgggagcgctgccccttcgaccgcagcctggccgt
gaccgtgccccaggccgcccgcgccgtggccgagggcagcgtggacctggaggacctggacgaggagcg
caagggctggagcggcgagcgcctgctggccgagctggagcgcacccgccccgccgacgaggacctggc

cgtgtgccacggccacctgtgccccgacaacgtgctgctggacccccgcacctgcgaggtgaccggcct
gatcgacgtgggccgcgtgggccgcgccgaccgccacagcgacctggccctggtgctgcgcgagctggc
ccacgaggaggacccctggttcggccccgagtgcagcgccgccttcctgcgcgagtacggccgcggctg
ggacggcgccgtgagcgaggagaagctggccttctaccgcctgctggacgagttcttctgagaattcgg
tac

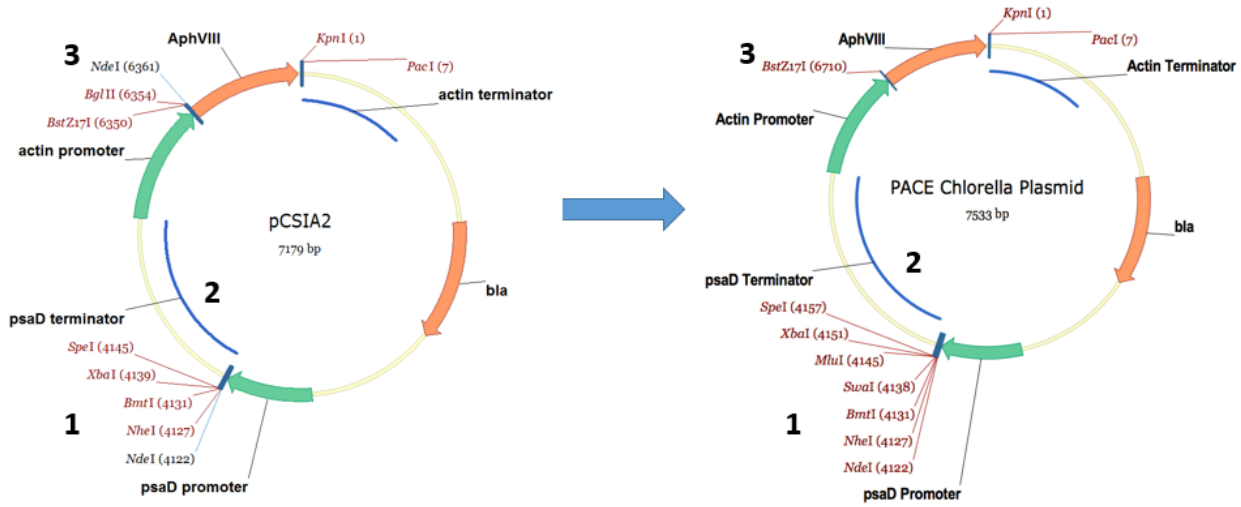


Items On Order

	Name	Product	Purification	Sequence
1	CSIprom1fw	25 nmole SameDay	Standard Desalting	GCG CTC TTC CGC TTT TCA GCA GGT TAT TGG
2	CSIMCS1fw	25 nmole SameDay	Standard Desalting	AGC CCC AGC CGA CCA GGA CAG AGG CCA AAC
3	CSIMCS1rv	25 nmole SameDay	Standard Desalting	GCA ACA AAA ACA GCA GCA GCA TTT AAT GAG
4	CSIMCS2fw	25 nmole SameDay	Standard Desalting	GGC ACC TTG CTG ACC TGA AGT TGG GGC GCG
5	CSIMCS2rv	25 nmole SameDay	Standard Desalting	CTC CCA GGT AGC AAC TAC TGC CGG CCC GAG
6	CSIterm2rv	25 nmole SameDay	Standard Desalting	GCG GCC GCA CAA TGC TGT GCT GCA GAG ACG
7	CSIprom2fw	25 nmole SameDay	Standard Desalting	ATC GAT GGC GGC AGC CAG CTC AAG TGT GTG
8	CSIterm1rv	25 nmole SameDay	Standard Desalting	ACC CTG CCC AGC ACT CCA GCC AGC AGC GAG

Figure 18: List of primers used by Genewiz to sequence confirm the pCSIA2 plasmid vector.

Figure 19: Sequence of 354 bp added the psaD promoter in pCSIA2 to create the 7533 bp PACE Chlorella Plasmid.



1. Expansion of MCS between psaD promoter and psaD terminator to add two restriction sites, SwaI and MluI by addition of 13 bp sequence, TAAATTCACGCGT, at position 4137. This resulted in the loss of the original **'C'** nucleotide at position 4137.
2. Expansion of psaD terminator sequence with the addition of the following sequence starting at nucleotide 4162 (numbering includes added sequence in #1):
   GCTGGCCACTCGGCCCAGCAGCTTGGCAGCGCTGGCCTGAGGGGCGCCTGGC
   CATGGCGTCCTGAGCGCTGGCGCAGCTTGCAGCCACCAGGGCCGCCGGTGCG
   GCGTGGCCGGCGGCCTGCCGGCCGCAGCTGCAGCTGCGGCGCGCGGCCGGCA
   GCGGCGGCGAAGGGGTGGTGGGCGGCGCTGCGGCGGCGTGGGGACGGCTCC
   CAGATTAGTGCCGTGCCCTCGTGCTGCGCCCGCCCAGCACTCAACCGCCGCCG
   CCCTCACCCACCCCCTGAGAAGAACTAACCAACCCTCACAGTTTCTAGTGCCA
   CAGCGTGTACGATACGTCCTGGCAATTATGTCTGT
3. Deletion of NdeI site at position 6361 of pCSIA2 resulting in a unique NdeI site at position 4122 in the final PACE Chlorella construct.
4. In total 361 bp were added and 7 bp were deleted from pCSIA2 ('C' at position 4137 and **CATATG** at position 6359-6364) resulting in a net gain of 354 bp to yield the 7533 bp PACE Chlorella Plasmid

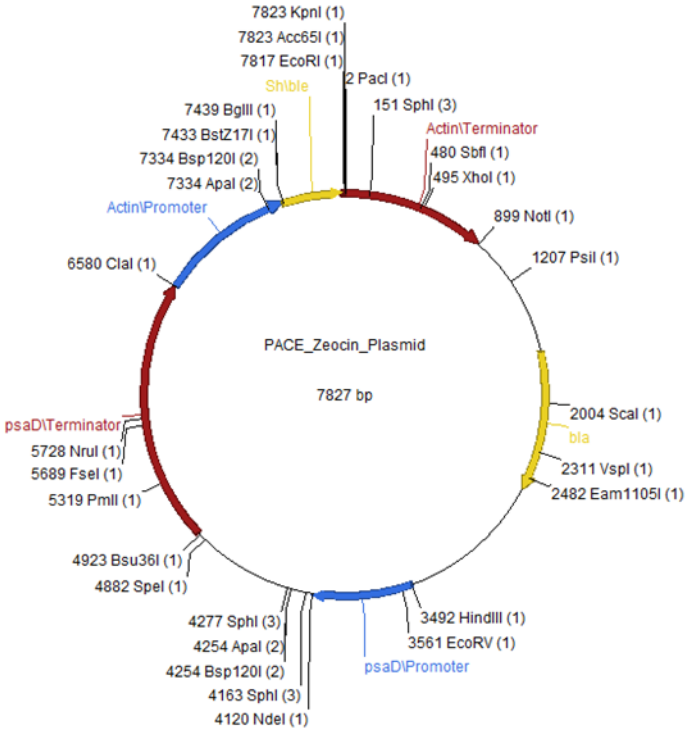


Figure 20: Map of PACE_Zeocin_Plasmid showing the Zeocin (Sh ble) gene insertion by replacing Paromycin (AphVIII) resistance gene

Figure 20A: Sequence of PACE_Zeocin_plasmid Sequence(map shown in Figure 20).

cttaattaagagctcctcgggccggcagtagttgctacctgggagtggtcggcgcggtgtgcacttgaagcaggggcggcttgcgtagtgcag
cagcgtgggaggaaattgcaaagactgtgcagcccacctcatctagcgcctgggatcgcatgcacacttcacatgctggcaatggtctaggc
cctgccctttcctggaggctgctctgccatacaccttgcctgacattacagcatggactgttgcatcgggtcgacagcgcacatgcaagccgag
ccacaccgctgcattcgagcatcaggccagcaccatgggctgcccaggcaggcagcctagcgtccgcccagcttctgccgcagggcttgca
gggcatccggcccgctggcctcccggcccagcgcagtggcaaccacagtcatgccacggccgcccgctttcttgcgctgcttggctgtcagg
gccttgttgagtgccgcctgcagggccgccgctcgagggtcaacaaccttcttttggcggccttcttcttctgctgcttggcgcccgatgagcccg
ctgtgccagctgcgcccgctgcactcgcccggctgcggctgtccgactgccgctgagccagagccacagcctggctcacagtctgcactgtgc
gaggctccggcacaggtgccgctgcctcgcctgccgctgccgcaccctctgctgctgccgccgctgctttcttgcgcttcttgccaccggctgctg
cgggggccgaagggtcgactgccgccatgactgctgctgctgcgccgccggtgaaccacttgcggctgaacttgggctgcagcggctcggc
cagcagcgctgccagctgtgcacgcacatctgctgccgtctctgcagcacagcattgtgcggccgccaccgcggtggagctccaattcgccct
atagtgagtcgtattacaattcactggccgtcgttttacaacgtcgtgactgggaaaaccctggcgttacccaacttaatcgccttgcagcacatc
ccccttcgccagctggcgtaatagcgaagaggcccgcaccgatcgcccttcccaacagttgcgcagcctgaatggcgaatggaaattgtaa
gcgttaatattttgttaaaattcgcgttaaatttttgttaaatcagctcatttttaaccaataggccgaaatcggcaaaatcccttataaatcaaaaga
atagaccgagatagggttgagtgttgttccagtttggaacaagagtccactattaaagaacgtggactccaacgtcaaagggcgaaaaaccg
tctatcagggcgatggcccactacgtgaaccatcaccctaatcaagttttttggggtcgaggtgccgtaaagcactaaatcggaaccctaaagg
gagcccccgatttagagcttgacggggaaagccggcgaacgtggcgagaaaggaagggaagaaagcgaaaggagcgggcgctaggg
cgctggcaagtgtagcggtcacgctgcgcgtaaccaccacacccgccgcgcttaatgcgccgctacagggcgcgtcaggtggcacttttcgg
ggaaatgtgcgcggaacccctatttgtttatttttctaaatacattcaaatatgtatccgctcatgagacaataaccctgataaatgcttcaataatatt
gaaaaaggaagagtatgagtattcaacatttccgtgtcgcccttattccctttttgcggcattttgccttcctgtttttgctcacccagaaacgctggtg
aaagtaaaagatgctgaagatcagttgggtgcacgagtgggttacatcgaactggatctcaacagcggtaagatccttgagagttttcgccccg
aagaacgttttccaatgatgagcacttttaaagttctgctatgtggcgcggtattatcccgtattgacgccgggcaagagcaactcggtcgccgca

tacactattctcagaatgacttggttgagtactcaccagtcacagaaaagcatcttacggatggcatgacagtaagagaattatgcagtgctgcc
ataaccatgagtgataacactgcggccaacttacttctgacaacgatcggaggaccgaaggagctaaccgcttttttgcacaacatgggggat
catgtaactcgccttgatcgttgggaaccggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctgtagcaatggcaa
caacgttgcgcaaactattaactggcgaactacttactctagcttcccggcaacaattaatagactggatggaggcggataaagttgcaggac
cacttctgcgctcggcccttccggctggctggtttattgctgataaatctggagccggtgagcgtgggtctcgcggtatcattgcagcactggggc
cagatggtaagccctcccgtatcgtagttatctacacgacggggagtcaggcaactatggatgaacgaaatagacagatcgctgagataggt
gcctcactgattaagcattggtaactgtcagaccaagtttactcatatatactttagattgatttaaaacttcatttttaatttaaaaggatctaggtgaa
gatccttttttgataatctcatgaccaaaatcccttaacgtgagttttcgttccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgag
atccttttttttctgcgcgtaatctgctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgccggatcaagagctaccaactctttttc
cgaaggtaactggcttcagcagagcgcagataccaaatactgttcttctagtgtagccgtagttaggccaccacttcaagaactctgtagcacc
gcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggcgataagtcgtgtcttaccgggttggactcaagacgatagttacc
ggataaggcgcagcggtcgggctgaacgggggggttcgtgcacacagcccagcttggagcgaacgacctacaccgaactgagatacctac
agcgtgagctatgagaaagcgccacgcttcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcggaacaggagagc
gcacgagggagcttccagggggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgatgctcgtca
ggggggcggagcctatggaaaaacgccagcaacgcggccttttacggttcctggccttttgctggccttttgctcacatgttctttcctgcgttatcc
cctgattctgtggataaccgtattaccgcctttgagtgagctgataccgctcgccgcagccgaacgaccgagcgcagcgagtcagtgagcga
ggaagcttttcagcaggttattggttcatgcacggtcatgttcctggtcgcttgtatctgcgatctgcaacgatatccatgatgacacgcaagtccac
gcattgagcagctccgtggtgccacagctggtccagccttgctacaagtgcgggtagagcaacagcacagccgtggcggctggaatttgcca
tggcctgtccatatctaccaaccagcgcgggatctgtgaagggcatgtgagcggcgccgcacgggcggcggctgccgccaaccggctgtcg
ccgcctgctgcagcgggtcgcctgacccccttgcagtgacttgctgttcacatcaagagcaatgaatgaatgaatgccctgccgcggtgaatgcccg
cagctgctgccccaggcagtcgggggacggcaccccagccggcaggtcgcccggcgcggattgccggcggctttctgaatgcaggcact
gccgataaagtggttgggctcccacacagcactttgcagcgccaccctgctgagcatttcaccgggtcacagcaggatcgtcagcgagaaag
gctcccccccagcgaccgctcggccgctgctcccacccccatatagccccagccgaccaggacagaggccaaaccatatgaccgtgaag
aagctgtacttcgtgcccgccggccgctgcatgctggaccactcctccgtgaactccaccctgacccccggcaacctgctgaacctgcccgtgt
ggtgctacctgctggagaccgaggagggccccatcctggtggacaccggcatgcccgagtccgccgtgaacaacgagaacctgttcgacg
gcaccttcgtggagggccagatcctgcccaagatgaccgaggaggaccgcatcgtgaacatcctgaagcgcgtgggctacgagcccgag
gacctgctgtacatcatctcctcccacctgcacttcgaccacgccggcggcaacggcgccttcaccaacacccccatcctggtgcagcgcgcc
gagtacgagaccgcccagcactccgaggagtacctgaaggagtgcatcctgcccaacctgaactacaagatcatcgagggcgactacga
ggtggtgcccggcgtgcagctgctgtacacccccggccacacccccggccaccagtccctgttcatcgagaccgagaactccggccccgtg
ctgctgaccatcgacgcctcctacaccaaggagaacttcgaggacgaggtgcccttcgccggcgtggactccgagctggccctgtcctccatc
aagcgcctgaaggagatcgtgcgcaaggagaaccccatcgtgttcttcggccacgacatcgagcaggagaagtcctgcaaggtgttccccg
agtacatctgatctagaactagtgctggccactcggcccagcagcttggcagcgctggcctgaggggcgcctggccatggcgtcctgagcgct
ggcgcagcttgcagccaccagggccgccggtgcggcgtggccggcggcctgccggccgcagctgcagctgcggcgcgcggccggcagc
ggcggcgaaggggtggtgggcggcgctgcggcggcgtggggacggctcccagattagtgccgtgccctcgtgctgcgcccgcccagcact
caaccgccgccgccctcacccacccccctgagaagaactaaccaaccctcacagtttctagtgccacagcgtgtacgatacgtcctggcaatt
atgtctgtctcattaaatgctgctgctgtttttgttgcaaaccgctagatttgcacacccctctctcccggccgtgtaaggctgaacccaacacgtgg
caccggggattgttgcgagcctggctggggaagacatgggcagctggttgagatgttgaattacagctgtacatcagtttgcgtgcattgtgtccg
aacctggagcactgcaaggcactggcggctggaagtggctggtcagttcactcttcattcaccagcaccacctcctcctgcaccagggcagcc
tcccctccgcagctgccgccgcagctgaacacagccacggtgagccactcccagctgtcgggcggctgcagggaggctgccgctgctgctc
cgccctcctcagctgccagccattcggctgcctcggtcaaagcggggatgaggggtgtcatcagctgcagctcaaacacgcgcggccggcc
gcaggcggcacagggtgcgggagcgggctgctcgcgagtcggccacagcggtgcaccgccaaagctgcagggtgaaggggccagagg
acagcagagaagtcggttttgggtcacttcagtagtcaccagcgggcagcaaaactgacctccagccctcttgcttgcactgaccacccctggg
cggccccgctcgcctgtagcgcgcgcactggtcggggcagcggcccagctgctttgcaaacttgaagaaggcggctcccacgcccgcccg
cgtgccgccggtgggcgccagcacagcgtcctcttcatagccttctcctgaccaggactcagggccgccctccacgctgggcgcacatgctgt
gcccgctgtggtcgccatggcggcgcccgcctcgccagctgctgccgcctctgcctcgtatcgagccatcagctgctgcaggtgctcctgttcgg

agcgctgctgcctgccgctgctggcgctgcactgcgactccggctcggcatacaggtagaaggcggggaggctgggctgggttgggtcatac
gctggcgctgccccagctgctgccacctcgctgctgcgcagctccccctgctgctgcggctgctgcttgccggctggcttggcagctggggcag
cactgcccaccgcttccagcgctgcgttgaggtctccaaagtcaaacggagctgccgcactgtccgtcacatcagcgcccatccccccagtcat
cgctgcctcctgcaccccagtcatctgcaccgcccataccccagtcatcggctgcaggcggcttctctgccggcggcgctgcctgctgctgcag
ctgctgcgctgtcggcggctgctcgctgctggctggagtgctgggcagggtatcgatggcggcagccagctcaagtgtgtgtggctctgggctg
ctgccggtgttgaccgccgctgcgccgccgcccggcagccagtgcaccagcagctgcgggccggcggacaggcacttgacatacagcgg
cgaacggctgcctgcaagcgggaatcggttgaggtcagctacggtgcctcgtgacggcagccagcagcgatcagcaagccacggctggg
agccatggcagccacccgaccaacacccaacaacccagcggcaacccaggtcagcagccgctgcacgcgccttgcgagtcttggtagcg
aaaggcgtaggaaccctccagagattcccacccagaggttcccacttcctcacggtcggcgccccagtctgcacatgtggatgggggtgggg
gcaattagtagcgaggatcggttaggcagggctggcacgcgagggcaaagggccgaacagggctcttggaaagctgatcatgcacccag
caagtgcaatcggccactgtgattggctgcttctctaaacatctgatcctgatccagcagacaaagccaagatccggcctctcaccggcaccct
cgtcctcggcggcagggccggcggccaccagcttgtagccttcggccagcaggaaggcgtgcaccgcaaacgccagcttgtcggcgtcgg
acctggcgagcatgagaagatacaacttggttgacgatgtgcaggtgtgctggactggccgaggggctggacggccgaggggcccgcagg
cggctggcggcgatggcggcggcgcggcggcggtgccagcagcagcgctttgcggcccggcaccttgctgacctgaagttggggcgcggt
atacagatctatggccaagctgacctccgccgtgcccgtgctgaccgcccgcgacgtggccggcgccgtggagttctggaccgaccgcctgg
gcttctcccgcgacttcgtggaggacgacttcgccggcgtggtgcgcgacgacgtgaccctgttcatctccgccgtgcaggaccaggtggtgc
ccgacaacaccctggcctgggtgtgggtgcgcggcctggacgagctgtacgccgagtggtccgaggtggtgtccaccaacttccgcgacgc
ctccggccccgccatgaccgagatcggcgagcagccctggggccgcgagttcgccctgcgcgaccccgccggcaactgcgtgcacttcgtg
gccgaggagcaggacgaattcggtac

B.3.b. Final recipient microorganism characterization

The confirmation of the randomly integrated DNA in the recipient strains was assessed by PCR-amplifying and sequencing the flanking regions (5' to the actin promoter for the ZeoR gene and 3' to psaD terminator for the CDS gene). Gene expression was assessed by using the RT-PCR and q-PCR analysis. Insert copy number will be determined by a Digital PCR assay currently being developed for green algae (https://www.thermofisher.com/us/en/home/life-science/pcr/digital-pcr/quantstudio-3d-digital-pcr-system.html).

B.3.c. Reference to any prior submission to EPA (or other Federal Agencies) which is directly related to construction of the subject microorganism (provide submission number if available).

We have not submitted this subject microorganism, nor do we have any knowledge of anyone else submitting this subject microorganism for EPA approval or to any other Federal Agency.

B.3.d. A key, which contains full names for abbreviations used in the diagram.
Abbreviations used in plasmid vector maps are listed in the legend of Figure 3 and in Tables 2 and 3.

B.3.e. As many circular plasmid/vector maps of intermediate constructs as necessary to clearly show genetic manipulations and gene modification. A linear portion of a plasmid representing only the changes is adequate for plasmids, which have been illustrated in their entirety earlier in the diagram. These intermediate construct illustrations should be sufficient to trace and verify the origins of intergeneric DNA shown in the final construct illustration.

See *section B.3.a.* above.

B.3.f. Sizes of important gene fragments retained and lost, sequences altered, and addition and/or deletion of restriction sites.

See Figure 2 and 3 for additions during vector construction. During the generation of the subject microorganisms, the paramomycin resistance gene (original selection marker in pSL18) was replaced by

the ZeoR resistance gene in the backbone vector pSL18. The actin promoter/terminator pair was introduced to drive the expression of the ZeoR gene.

B.3.g. Restriction enzymes used, including whether resulting cuts are full or partial.

BglII and EcoRI were used by Genewiz to clone SH *ble* between the actin promoter/terminator pair. NdeI and SpeI were used to clone SNRK 2 gene between the psaD promoter/terminator pair. ScaI was used to linearize the plasmid before particle bombardment transformation was performed.

B.3.h. Detailed map of the final cloning vector if the vector, or portions of that vector, are to be retained in the subject microorganism.

The critical components to integrate into the nuclear genome are the psaD promoter sequence, integrating gene coding region (SNRK 2), psaD terminator, actin promoter, selection marker (ZeoR) and actin terminator. These regions are located between the NotI and HindIII restriction enzyme sites on each plasmid as depicted in Figure 21.

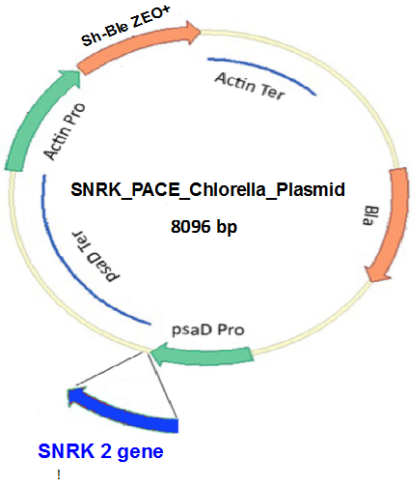


Figure 21: Plasmid maps of the final construct used to transform the recipient microorganism, *C. sorokiniana* 1412, and create the subject microorganisms, PACE_CS1412_ SNRK 2. The integrating region of gene in plasmid starts at the *HindIII* restriction enzyme site at neucleotide 3492 on each plasmid and encompasses the entire sequence through the NotI site at neucleotide 899.

B.3.i. Methods for isolating and identifying the DNA used to modify the recipient microorganism.

Genomic DNA was isolated from the recipient microorganism using Quick-DNA Fungal/Bacterial Kit from Zymo research (https://www.zymoresearch.com/dna/microbial-environmental-dna-isolation-1/bacterial-fungal-dna/quick-dna-fungal-bacterial-miniprep-kit). The presence of the gene integration was assessed via PCR amplification of genomic DNA (see Part 5 of section *B.3.m;*) using the primers listed in Table 3, which were specifically designed to cover the region from the promoter to the gene of interest. The PCR fragments were then confirmed by sequencing.

Table 3: Primer sequences used to identify the genomic DNA incorporated into the recipient organism.

Plasmid	Primers
SNRK_PACE_Chlorella_Plasmid	Fwd: TCCTGGTCGCTTGTATCTGC Rev: TAAGCGGAAGATTGCTCGGG

B.3.j. Preparation or modification of DNA including procedures such as deletion, insertion, directed mutagenesis, and rearrangement

There were no alterations to the DNA described above.

B.3.k. Procedures for selection of intermediate hosts including methods and results for determining the success of insertion, deletion, and/or rearrangement

No intermediate hosts other than *E. coli* for replication of the plasmids. All constructs were sequenced to confirm their integrity.

B.3.l. Characterization of vectors so that their function (cloning, expression, or shuttle) is noted.

All vectors were sequenced and identity subsequently confirmed by diagnostic restriction enzyme digest patterns on gels. The vectors are linearized to allow for non-homologous random DNA insertion into the plant chromosome. They cannot replicate autonomously in the algae. They do contain algal nuclear gene promoter/terminators (see maps) to drive the expression of the gene of interest. The characterization of the vectors used in this project is listed in Table 4.

Table 4: Characterization of the vectors described in this TERA submission

Vector	Function
SNRK_PACE_Chlorella_Plasmid	Expression

B.3.m. Transfer or integration techniques such as transformation, filter matings, triparental matings, and recombination events critical to the construction of the subject microorganism.

Below is the protocol used to introduce DNA into the recipient microorganism.

Part 1: Algal culture preparation:

Chlorella sorokiniana (strain 1412) cells are grown in 100 mL of HS media on a shaker platform at 120 rpm at a light intensity 50 µmol m^{-2} s^{-1}. The culture was harvested after reaching an OD_{750} 0.3-0.5 (exponential phase). A 50 ml aliquot at OD_{750}= 0.3-0.5 was spun down at 3000 rpm for 5 min on a tabletop centrifuge.

Part 2: Electroporation process:

Following steps were performed on ice.

The cell pellet was washed two times with sterile 375 mM cold sorbitol and resuspended in 5 ml of 375 mM cold sorbitol. The cell suspension was kept on ice until electroporation. For transformation, 250 µL of the cell suspension was mixed with ~ 1 µg linearized plasmid DNA and placed it in a pre-chilled 4 mM gap electroporation cuvette. The DNA/cell mix was pulsed three times (~10 s between each pulse) in a Biorad

(Hercules, CA) Gene Pulser Xcell using the program setting "time constant" with the following settings: 20 ms pulse, 1600 V. Electroporated cells were incubated on ice for 5 minutes and then resuspended in 5 ml of TAP media in a culture tube.

Part 3: Screening of transgenics:

250uL of culture was then spread on each large selection plate (PROTEOSE + ZEOCIN 40ug/mL +AMP 50ug/mL). Cells numbers may vary from experiment to experiment.
Typically, 15 - 30 antibiotic resistant colonies are selected for further analysis for PCR confirmation of gene of interest.

Part 4: Best practices (required):

1. At least ten PCR positive colonies should be screened to obtain a broad range of phenotypes
2. Gene of interest should be confirmed in transgenic algae by both PCR and DNA sequence analysis, particularly if the gene of interest is mutated.
3. Ideally, transformants should be single copy insertions to avoid gene silencing. This can be confirmed by Southern blotting or qPCR.
4. Expression levels of gene of interest should be evaluated by qPCR versus housekeeping standard or by western blot analysis in a linear range of antibody response.
5. Maintain transformants either on antibiotic selection plates or by freezing. The transgenics should be tested occasionally for expression of gene of interest and results recorded.
6. A uniform coding system shall be used for all transformants based on the following coding system: PACE: strain-gene of interest name- transformation event number.
7. Records of transformants will be kept in an Excel spreadsheet with all relevant information listed in 6 above plus date of transformation and name of person who generated transformants.
8. Transformants should also be bar-coded (will get info on bar coders) and transferred individually (one transformant in hood at a time) in sterile transfer hoods when needed to avoid miss labeling.
9. An Excel spreadsheet shall be generated to track transfers of individual transformants to new media including; transformant name, date transferred, media used, person's name.
10. When transgenics are shipped to other labs, PCR DNA primer information should be included for the gene of interest and the recipient lab should confirm transgene presence by PCR. In addition, records of transfer events for each line shall be sent electronically.

Part 5: Screening of putative transformants:

1. DNA extraction of colonies: after colony selection on proteose plates containing zeocin 40ug/mL, a speck of algae is suspended in 50uL of Chelex 5% v/v, mixed thoroughly, heated at 98C for 15 min, and then centrifuged at 10g to get rid of cell debris.
2. 3uL of the supernatant is used as template for PCR reactions:
    a. PCR conditions using NEB Q5 kit:

1) 98C 30 sec 2) 98C 10 sec, 68C-1 annealing, and 72C 35sec for elongation x4 3) 98C 10 sec, 63C annealing, and 72C 35 sec for elongation X30

    b. Primers:

PCR primers:
Fwd: TCCTGGTCGCTTGTATCTGC
Rev: TAAGCGGAAGATTGCTCGGG

B.3.n. Catalogue references for commercial systems used such as special recipient strains, plasmids, cosmids, etc.

The source and identity of the *E. coli* plasmid used to clone the final constructs is noted in the reference below. *Sh* ble, the gene that confers resistance to Zeocin, is from *Streptoalloteichus hindustanus.*

https://www.thermofisher.com/us/en/home/life-science/cloning/competent-cells-for-transformation/competent-cells-strains/dh5a-competent-cells.html?gclid=CIiV0LTJ6tICFQSOaQodSWQKiA&s_kwcid=AL!3652!3!163711299069!e!!g!!%2Bdh5%20%2Balpha&mkwid=sy49S36eD-dc_pcrid_163711299069_pkw_+dh5%20+alpha_pmt_e_slid__dimid=&ef_id=WKj4ewAAAD02dRYn:20170322165239:s

B.3.o. Literature references to original sources of important sequences used in both intermediate and final constructs

Table 5: References to sources of intergeneric sequences in final constructs.

Sequence	Reference
SNRK 2	Gonzalez et al., 2018

B.4 Properties of the Subject Microorganism

B.4.a. Methods and results used to verify the final construct, including verification of the location of the intergeneric DNA and its copy number

The confirmation of the randomly integrated DNA in the recipient strains was assessed by PCR-amplifying and sequencing the flanking regions (regulatory elements) and SNRK 2 gene (Figure 22) and expression of the SNRK2 gene was validated by RT-PCR and q-PCR analysis (Figure 23).

Insert copy number will be determined by a Digital PCR assay currently being developed for green algae (https://www.thermofisher.com/us/en/home/life-science/pcr/digital-pcr/quantstudio-3d-digital-pcr-system.html).

Figure 22: Confirmation gel of the randomly integrated DNA in the recipient strains

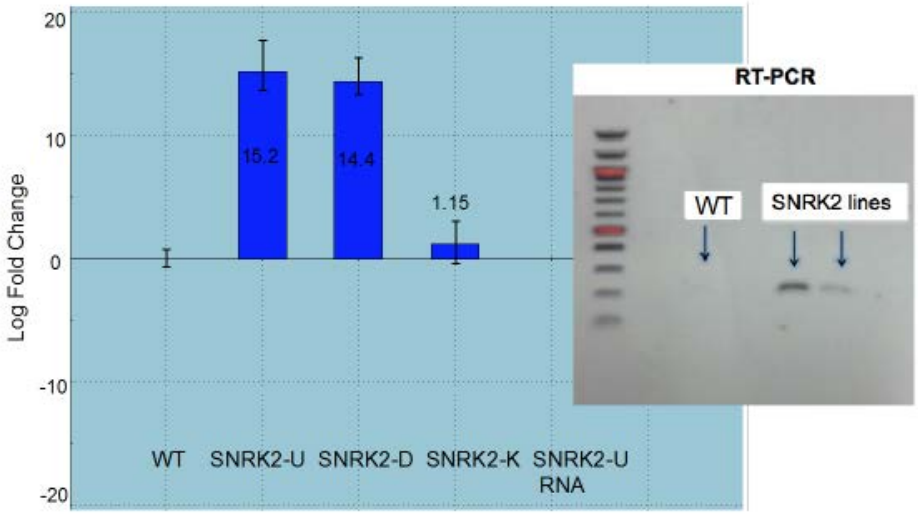


Figure 23: Expression of the SNRK2 gene validated by RT-PCR and q-PCR analysis

B.4.b. Nucleotide and protein sequences of major structural gene(s), if appropriate.

The gene included in the transformation constructs SNRK. This gene was expressed in its native form in the recipient organism, synthesized by a contract vendor directly into the PACE Chlorella plasmid described in section B1.c. The nucleotide and protein sequences for these genes are shown below:

SNRK 2 Nucleotide Sequence:

ATGGATATGCTACCTGAACGTTACGAGCTGATAAGAGACATTGGCAGTGGAAATTTTGGTGTAGCCAAGTTAATGCGTGA
CCGAAAAACCGGCGAGCAGGTTGCTGTGAAGTTCATCGAGAGAGGAGAAAAGATCGATAAGAATGTCGAGCGTGAAAT
TGTCAATCACCGTCAACTTTCTGGACACCCCAACATAATACGGTTCAACGAGGTCTTTCTCACCTCAACTCATTTAGGCATA
TCTATGGAATATGCTTCAGGAGGAGAATTATTCGATCGTATCGTAAAGGCTGGGAGGTTTTCAGAGGATGAGGCTAGATA
TTTCTTTCAGCAGTTGATATCAGGAGTAAACTGGTGCCACAAAGAGGGTGTATGCCATAGGGACCTAAAATTGGAAAATA
CGCTACTAGATGGTAGAGCTGCACCAAGATTGAAAATATGCGATTTTGGATATTCTAAATCGGCTATTTTTGACTCGCAGC
CGAAATCTACCGTAGGTACGCCCGCCTATATAGCGCCTGAAGTTCTATCCAGAAAGCAATATGCGGGGGAGATTGCTGAT
GTTTGGTCATGTGGGGTAACCTTGTACGTTATGCTAGTGGGTGCATATCCCTTTGAAGATGCAAACGATCCTAGAAACTTT
AGAAAAACCATCCAACGGATCATGGGTGTGAAATACTCGTTCCCGAGCAATCTTCCGCTTAGTAAAGAGTGTCATGATCT
GATGAGTAGGATTTTTGTTGCTAATCCATCCAAAAGAATTTCATTGACGGAGATTAAGGCACATCCATGGTTTTTGAAGAA
CTTGCCTCCAGAATTAGATGAGAAAGAAATTATCGGAGAAGTGAGATCAAAAAACTCTGATATACAGCCAGTAGAGGAG
ATCAAGAGACTTGTTGCAGATGCAAGGATGAGACCAAGTCAAGCTGGAATGCATACCTTCAACGAAGATGACTACATGG
ACGGCGATATAGATATGGAGACAGAAATTATGGATTAA

SNRK2 Protein Sequence:

MASASMDMLPERYELIRDIGSGNFGVAKLMRDRKTGEQVAVKFIERGEKIDKNVEREIV
NHRQLSGHPNIIRFNEVFLTSTHLGISMEYASGGELFDRIVKAGRFSEDEARYFFQQLIS
GVNWCHKEGVCHRDLKLENTLLDGRAAPRLKICDFGYSKSAIFDSQPKSTVGTPAYIAPE
VLSRKQYAGEIADVWSCGVTLYVMLVGAYPFEDANDPRNFRKTIQRIMGVKYSFPSNLPL
SKECHDLMSRIFVANPSKRISLTEIKAHPWFLKNLPPELDEKEIIGEVRSKNSDIQPVEE
IKRLVADARMRPSQAGMHTFNEDDYMDGDIDMETEIMD

ZeoR Nucleotide Sequence:

ATGGCCAAGCTGACCTCCGCCGTGCCCGTGCTGACCGCCCGCGACGTGGCCGGCGCCGTGGAGTTCTGGACCGACC
GCCTGGGCTTCTCCCGCGACTTCGTGGAGGACGACTTCGCCGGCGTGGTGCGCGACGACGTGACCCTGTTCATCTCC
GCCGTGCAGGACCAGGTGGTGCCCGACAACACCCTGGCCTGGGTGTGGGTGCGCGGCCTGGACGAGCTGTACGCC
GAGTGGTCCGAGGTGGTGTCCACCAACTTCCGCGACGCCTCCGGCCCCGCCATGACCGAGATCGGCGAGCAGCCCT
GGGGCCGCGAGTTCGCCCTGCGCGACCCCGCCGGCAACTGCGTGCACTTCGTGGCCGAGGAGCAGGAC

ZeoR Protein Sequence:

MAKLTSAVPV LTARDVAGAV EFWTDRLGFS RDFVEDDFAG VVRDDVTLFI SAVQDQVVPD
NTLAWVWVRG LDELYAEWSE VVSTNFRDAS GPAMTEIGEQ PWGREFALRD PAGNCVHFVA
EEQD

B.4.c. Description of gene regulation and expression in the subject microorganism and the characteristics of the product encoded for the intended use.

SNRK 2 gene is shown to play a critical role in Arabidopsis plants to regulate the energy metabolism. The overexpression of SNRK 2 gene in Arabidopsis increased sucrose synthesis, starch synthesis and leaf growth (Zheng et al 2010). We expected overexpression of SNRK 2 gene would improve starch accumulation and growth in Chlorella cells (See section *B.4.f.)*. The regulatory elements used to express SNRK 2 gene is the psaD (a photosynthesis-related genes) promoter and terminator, which are endogenous to the recipient microorganism.

B.4.d. Possibility of unexpected gene expression or suppression in the subject microorganism as a result of the genetic modification, both during use and subsequent to release to the environment.

There is no reason to expect that changes in gene expression within the subject microorganism will result from the genetic modifications described in this document.

B.4.e. Indication of whether the major gene product is extracellular or intracellular.

SNRK 2 is expressed in the cytosol, therefore intracellular proteins.

B.4.f. Methods and results for determining stability of the introduced DNA

The stability of the DNA can be assessed by PCR and qPCR after multiple generations of growth. We have been monitoring the stability of gene insertion through both PCR and q PCR. We have included the data in section *B.4.a.* We have further performed the phenotypic characterization of positive transgenics to determine the result of gene insertion. The incorporation of SNRK 2 gene resulted in better growth and photosynthetic efficiency.

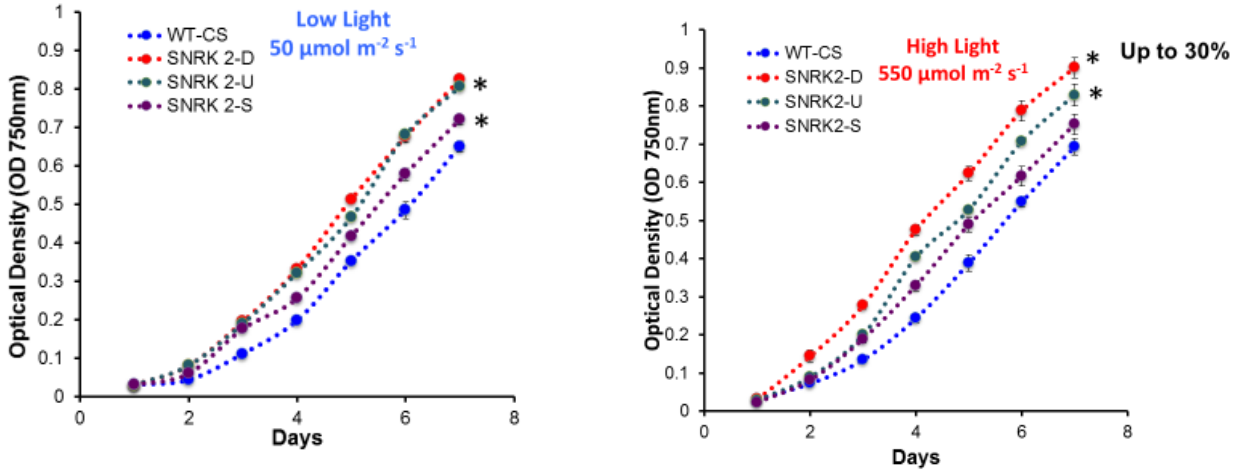


Figure 24: SNRK 2 overexpression leads to improved growth in both low and high light intensities. Experiments were performed in Triplicates. * Indicates p-value < 0.005

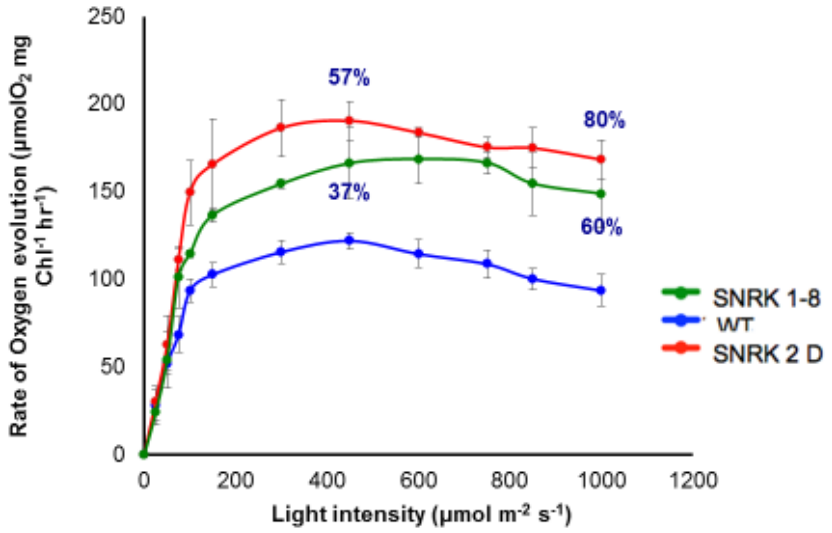


Figure 25: SNRK 2 overexpression leads to improved photosynthetic efficiency. Experiments were performed in Triplicates. Cells were grown in Shaker Flasks for 3 days. Light intensity 300 µmol m^{-2} s^{-1}

B.4.g. Likelihood of genetic transfer of the introduced DNA by transformation, conjugation, transduction, and/or transfection during use and subsequent to release to the environment.

As evidenced by the stability of the DNA (discussed in Section B4.f) and as discussed in Section D2.f, there is the possibility for gene transfer through mating with wild *C. sorokiniana* strains.

B.4.h. Growth characteristics in laboratory and environmental media such as generation time, growth temperature (optimum and range), pH (optimum and range), oxygen requirements (optimum and range), preferred energy and carbon sources, etc.

C. sorokiniana (DOE1412) cells are grown autotrophically in flasks in inorganic media on a shaker platform at 120 rpm at a light intensity ~ 50 uE/m^2/sec PAR and *pH* ~7. No carbon source is added to the media. In general, *C. sorokiniana* (DOE1412) grows well at temperatures from ~15 °C – 32°C with an optimum around 28 °C and can tolerate short duration excursions to higher temperatures (up to 40 °C, although with much reduced growth rate). It can grow in range of salinity from freshwater to nearly 1x seawater, though it has to be adapted onto higher salinities and does not tolerate large swings in salinity well. It will grow on a range of nutrient sources. For specific scale-up culture conditions for subject and recipient microorganism under this experiment plan see section *F.1.a.3.*

B.4.i. Factors limiting growth, survival, or reproduction (such as auxotrophy, asporogenicity, debilitation from continuous culturing, etc.

No factors have been introduced that limit growth.

B.4.j. Subject microorganism antibiotic profile.

During the course of the field experiment, it will be necessary to inactivate the subject microorganism. We will utilize treatment of biomass by bleaching and/or autoclaving as a way to inactivate cultures and biomass. The engineered algae are zeocin resistant but this antibiotic will not be used in the field under any circumstances.

C. POTENTIAL HUMAN HEALTH EFFECTS OF THE SUBJECT MICROORGANISM

C.1 Pathogenicity of Subject Microorganism

C.1.a. Nature and degree of pathogenicity, virulence, or infectivity, in humans.

The subject microorganism, which is taxonomically the same as its parental strain, *C. sorokiniana,* is not likely to pose any risk to human health. There is no evidence in the published literature that *C. sorokiniana,* is pathogenic, virulent or infective in humans and no evidence to suggest that it has any pathogenic, infective or toxic properties. *Chlorella sp.* are generally regarded as safe (GRAS) for human consumption according to the FDA (Bagchi 2016; Roberts 2016). Literature searches using the terms Chlorella* toxic* pathogen* infect* virus* in PubMed and Web of Science returned no articles demonstrating that *C. sorokiniana* is harmful to humans. A review of EPA's Drinking Water Contaminant Candidate List 3 (CCL3) and their List of Pests of Public Health Importance (Pesticide Registration Notice 2001-1) did not include any *Chlorella sp.* The genus *Chlorella* was also not found on the Risk Group Database of the American Biological Safety Association (https://my.absa.org/tiki-index.php?page=Riskgroups. In addition, a review of the CDC websites did not associate *Chlorella sp.* with any adverse health effects and the American Type Culture Collection (ATCC) lists *C. sorokiniana,* along with 22 other strains of the genus *Chlorella,* as Biosafety Level (BL1) 1 organisms based upon the fact that these organisms are not known to cause disease in healthy humans.

C.1.b. Results of pathological tests on effects of the subject microorganism in mammalian species, if relevant and available.

There are no reports of *C. sorokiniana* having a pathogenic effect in mammalian species. There are rare instances of *Chlorella sp.* infections in animals, primarily immunocompromised sheep drinking from contaminated water sources (Ramirez-Romero, 2010). In contrast, *Chlorella sp.* has been used as a supplement to animal feed and has been proposed as a protein supplement for human consumption (Becker 2007). In humans *Chlorella sp.* supplements have shown beneficial effects including improved immune responses, improved healing of the small intestine epithelium, antioxidant action and even antitumoral effects (Ramirez-Romero, 2010).

C.1.c. Ability to colonize humans (e.g., the skin, the gut, etc.).

There is no evidence in the literature that *C. sorokiniana* has the ability to colonize humans. See references in section *C.1.a.*

C.1.d. Ability to grow at human body temperature, 37° C.

As discussed in section B4.h, the recipient microorganism has the ability to grow at 37° C, but due to its photosynthetic nature and other requirements for growth it is highly unlikely that the organism could survive and replicate inside the mammalian body.

C.1.e. Susceptibility to control measures such as antibiotics or disinfectants, substrate requirements, or physical and chemical control methods.

The subject microorganism is susceptible to control measures, such as a 5% bleach solution, autoclaving at 120° C for 60 minutes.

C.2 Toxicity and Immunological Effects of Subject Microorganism or Its Products

C.2.a. Nature and degree of toxicity to humans.

As described in Section C1.a, the subject microorganism is not expected to have any toxic effect on humans. There have been no studies that indicate that any of the members of the *Chlorella* species synthesize or secrete toxins. Toxic microalgae are common among the cyanobacteria and dinoflagellates (Metting and Pyne 1986), which are unrelated to the recipient microorganism, *C. sorokiniana*.

C.2.b. Results of toxicological tests on effects of the subject microorganism or its products in mammalian species, such as allergenicity, or immunological responses after exposure via ingestion, inhalation, or dermal contact.

As described in section *C.1.a,* the subject microorganism, which is taxonomically the same as its parental strain, *C. sorokiniana,* is not likely to pose any risk to human health. There is no evidence in the published literature that *C. sorokiniana* is pathogenic, virulent or infective in humans and no evidence to suggest that it has any pathogenic, infective or toxic properties. *Chlorella sp.* are generally regarded as safe (GRAS) for human consumption according to the FDA (Bagchi 2016; Roberts 2016). *Chlorella* is a popular human nutritional supplement and extracts are used in skin care products.

A FASTA search against the allergenonline.com database (v18B, March 23, 2018) was conducted on the ZeoR and SNRK2 amino acid and protein sequences. No results were yielded for ZeoR nor were any results yielded with a sliding window 80-mer or exact match 8-mer. For SNRK2, the ORF's (identified using ExPasy at https://www.expasy.org/) yielded the FASTA search results below in Table 6 with no hits above 35% identity for the 80-mer sliding window and the 8-mer exact match searches yielded no hits.

Table 6: FASTA search results for SNRK2 ORF

ORF	Amino Acid Sequence	AllergenOnline Search Results		
		FASTA	80-mer	8-mer
5'3' Frame 1	MASASMDMLPERYELIRDIGSGNFGVAKLMRDRKTGEQV AVKFIERGEKIDKNVEREIVNHRQLSGHPNIIRFNEVFLTST HLGISMEYASGGELFDRIVKAGRFSEDEARYFFQQLISGV NWCHKEGVCHRDLKLENTLLDGRAAPRLKICDFGYSKSAI FDSQPKSTVGTPAYIAPEVLSRKQYAGEIADVWSCGVTLY VMLVGAYPFEDANDPRNFRKTIQRIMGVKYSFPSNLPLSK ECHDLMSRIFVANPSKRISLTEIKAHPWFLKNLPPELDEKEI IGEVRSKNSDIQPVEEIKRLVADARMRPSQAGMHTFNEDD YMDGDIDMETEIMD**STOP**	No Matches	No Matches of greater than 35% identity	No Matches
5'3' Frame 2	MSSVKLSITVNFLDTPT**STOP**	No Matches	No Matches of greater than 35% identity	No Matches
5'3' Frame 3	MWGNLVRYASGCISL**STOP**	No Matches	No Matches of greater than 35% identity	No Matches
3'5' Frame 1	MIRWMVFLKFLGSFASSKGYAPTSITYKVTPHDQT SAISPAYCFLDRTSGAI**STOP**	1 Match at 64.4%	No Matches of greater than 35% identity	No Matches
3'5' Frame 1	MTKHQQSPPHIAFWIELQALYRRAYLR**STOP**	No Matches	No Matches of greater than 35% identity	No Matches
3'5' Frame 1	MTLFTKRKIARERVFHTHDPLDGFSKVSRIVCIFK GICTH**STOP**	1 match at 60.0%	No Matches of greater than 35% identity	No Matches

C. sorokiniana has been cultivated by members of the consortium (e.g., AZCATI/LANL/Reliance) for several years, both in closed reactors and outdoor ponds. Throughout this time no employees have reportedly suffered any negative consequences by contact with this strain. Considering the ubiquitous nature of microalgae and the *Chlorella* genus, it is highly likely that most humans have been exposed to this species already with no evidence of harm. Workers will have minimal exposure to the subject microorganism and will be required to wear the appropriate Personal Protective Equipment (PPE) throughout the proposed experimental plan. Direct contact with the subject microorganism will only occur for brief periods of time for sampling and analysis. Taken together, neither the recipient algae nor the introduced traits have toxic or allergenic qualities that should raise any concern for the subject microorganism.

D. PREDICTED ENVIRONMENTAL EFFECTS AND FATE OF THE SUBJECT MICROORGANISM

D.1 Ecological Effects

D.1.a. Nature and degree of pathogenicity, virulence, or infectivity to mammals, fish, insects and other invertebrates, and plants; including host range.

The subject microorganism is of the genus *Chlorella*. Species within this genus have been found as part of the freshwater plankton in rivers, ponds, and lakes, as well as in marine and edaphic habitats around the world (Bock 2011; Ratha 2012). Like many other algae, *Chlorella* plays an important role in the aquatic food chain as a primary producer and food source for higher trophic levels.

Algae are known to be sources of many useful products such as vitamins, amino acids, fatty acids, and simple carbohydrates that are essential or support the growth of other microbes (Metting and Pyne 1986). Although these compounds have biological activity, none from *C. sorokiniana* have been identified or strongly implicated as producing antibiotics, algicides, toxins, pharmaceutically active compounds, or plant growth regulators. Extracts of *C. sorokiniana* have been reported to have anti-bacterial properties when tested against the human pathogenic strains *Escherichia coli*, and *Staphylococcus aureus* (Li et. al. 2016). However, it should be noted that this was a comparative study of whole cell extracts of microalgal strains from the Hainan province of China and the specific agent conferring the antibacterial properties was not isolated hence, other factors such as extraction solvent, temperature of incubation, pH of culture medium, incubation period, medium constituents and light intensity could have influenced the antimicrobial activity detected.

As described in Section C, there is no evidence to suggest that *C. sorokiniana* has any harmful properties to humans. In particular, it is not known to be pathogenic, virulent or infectious to humans or animals, nor does it have any deleterious effects on plants. Literature searches on PubMed and Web of Science using the terms Chlorella* toxic* pathogen* infect* virus* do not return any articles that indicate that *C. sorokiniana* has the potential to be harmful. Literature searches and sequence analyses have also failed to identify any evidence of any such harmful properties.

D.1.b. Toxicity of microbially-produced toxins to mammals, fish, insects and other invertebrates, and plants.

As described in detail in Section C, *C. sorokiniana* is not known to produce any toxins that might be harmful to humans, animals, or plants. Algal toxins are best known from dinoflagellates and cyanobacteria, and a few diatoms, haptophytes and euglenoids; none of which are related to *C. sorokiniana.*

D.1.c. Involvement in or effects on biogeochemical processes (e.g., effects on nutrient cycling, particularly C,N,P, and S; effects on primary production (CO2 fixation); utilization of complex carbon substrates, such as cellulose and lignin degradation; effects on nitrogen fixation; effects on nitrification.

Chlorella sp., as with many other microalgae, are an important primary producer and food source for higher trophic levels. In addition to the four basic elements; carbon, nitrogen, phosphorus, and sulfur, ionic components such as sodium, potassium, iron, magnesium, calcium, and trace elements must also be provided for algal growth. Carbon, in the form of carbon dioxide, may be fixed from the atmosphere through photosynthetic activity. Organic forms of carbon can also be used when the algae possesses both autotrophic and heterotrophic traits. Carbon can also be utilized in the form of soluble carbonates for cell growth, either by direct uptake or conversion of carbonate to free carbon dioxide through carbonic anhydrase activity. Microalgae play a key role in converting inorganic nitrogen to its organic form through a process called assimilation. Inorganic nitrogen can be in the forms of nitrate, nitrite, and ammonium. All forms of inorganic nitrogen are ultimately reduced to ammonium prior to being incorporated into amino acids within the intracellular fluid. Inorganic phosphates play a significant role in algae cell growth and metabolism as well. All of these are readily found in many environments.

The modifications made to the recipient microorganism in this TERA in clove the introduction of the SNRK 2 gene from *Picochlorum soloecismus* strain, which is a genus of green algae in the class Trebouxiophyceae. The modification is not expected to introduce any other phenotypic change in the recipient microorganism, and will not impart or enhance any harmful traits beyond what may be present in the recipient strain.

D.1.d. Known or predicted effects on other organisms including microorganisms in the environment, including effects on competitors, prey, hosts, symbionts, predators, parasites, pathogens; effects on community structure or species diversity.

C. sorokiniana is a primary producer and does not require a host organism to survive. No species have been identified as prey, hosts, symbionts, or parasites to *Chlorella*. We do not expect that the genetic modifications introduced into the subject microorganism will impart any traits that could lead to any adverse effect on any other organisms in local environments in which the subject microorganism would be used. Pathogens with the ability to impair and kill microalgae include: viruses, bacteria, fungi, and a number of protists. Primary predators that have been observed at the AzCATI site for *Chlorella sp.,* including the recipient microorganism to be used, have included *Vampirovibrio chlorellavorus*, a gram-negative obligate aerobic and epibiotic parasitic bacterium (Soo et al. 2015) and has been shown to bring down *Chlorella sp.* cultures at AzCATI and other sites in particular in the southwest very rapidly. In addition, *Poterioochromonas sp*., a golden algae that is flagellated, is a particular persistent predator at the AzCATI site and has significantly limited cultivation of *C. vulgaris, C sorokiniana,* and *C. zofingiensis.* In salt water adapted cultures of *C. sorokiniana*, we have seen invasion by amoeba. Other zooplankton "grazers", such as rotifers and ciliates are also present and can be a common source of culture collapse. While is it not expected that the subject microorganisms will have any impact on competitors or community structure, we will test their ability to do so in our experimental plan.

D.1.e. Identification and description of target organisms, e.g. taxonomy, agricultural uses, and the anticipated mechanism of interaction between microorganism and target organism.

C. sorokiniana has no 'target organism'. No species have been identified as prey, hosts, symbionts, or parasites. Additionally, no hits were returned when the search terms **C. sorokiniana* and *target organism were used in a PubMed search. The genetic modifications introduced into the subject

microorganism are not expected to impart any traits that could lead to any adverse effect on any other organisms in local environments in which the subject microorganism would be used.

D.1.f. Existence of nontarget organisms or alternate hosts (e.g., nitrogen fixing bacterial inoculants often have an intended legume host, but may be able to infect other leguminous plants).

The microorganism is not designed to be a host or to infect or feed upon any other living organisms. No host is needed for *C. sorokiniana* to survive, as it is a primary level producer in aquatic ecosystems serving in the fixation of inorganic substrates and providing nutrients to higher trophic levels. Additionally, no hits were returned when the search terms **C. sorokiniana* and *non-target organism or *alternative host were used in a PubMed search. *C. sorokiniana* is likely to mate with other wild *C. sorokiniana* unless mating is impaired, for example, by doubling chromosomes numbers using colchicine, photoropin mutants, or the use of inducible terminator gene technologies which are under development but not part of this experiment.

D.1.g. Known or expected substrate range of degradative gene protein products including both contaminant compounds and environmental substrates (e.g., lignin).

Neither the recipient, nor subject microorganism contain (or are expected to contain) any degradative gene protein products. No hits were returned when the search terms **C. sorokiniana* and *degradative gene product were used in a PubMed search. Additionally, no hits were returned when *pyrroline-5-carboxylate synthase or AHL-Lactonase and *degradative gene product were used in a PubMed search.

D.1.h. Known or expected metabolic pathways of xenobiotic contaminant(s) present.

D.1.i. Nature and degree of toxicity of metabolites (dead-end or intermediate metabolites produced by metabolism of a xenobiotic contaminant) to mammals, fish, insects and other invertebrates, and plants. Toxicity should be compared with the toxicity of the parent contaminant compound.

D.1.j. Resident antibiotic production levels.

1h-j. Neither the recipient nor the subject microorganism produce or contain any xenobiotic contaminants. A PubMed search using the terms **C. Sorokinana* and *xenobiotic retrieved 1 journal article which described the effects of xenobiotics on two Chlorella species. There are no reports that *C. sorokiniana* produces any antibiotics. When the search terms **C. sorokiniana* and *antibiotic production were used in a PubMed search only one hit was returned (Tishchner 2004). The article, titled: "Mitochondrial electron transport as a source for nitric oxide in the unicellular green alga *Chlorella sorokiniana"* did not mention, refer to, or discuss, antibiotic production by *C. Sorokiniana.*

D.2 Survival and Fate

D.2.a. Natural habitats and geographical distribution of the recipient microorganism.

Strains of *C. sorokiniana* belong to the Order *Chlorellales* in the Class *Trebouxiophyceae. Chlorella sp.* are commonly found in the plankton of freshwater rivers, ponds, and lakes, and sometimes in brackish or marine habitats. The genus has been found throughout all of North America from tropical to arctic climates. Reports in the literature indicate that *Chlorella sorokiniana* is likely to be global in distribution including India (Ratha et al. 2012). These organisms use light energy to convert CO^2 and H_2O into carbohydrates and other cellular products. Growth may be dense in nutrient-rich waters but is not typically considered a nuisance. Like many other algae, *Chlorella* is an important primary producer and food source for higher trophic levels. *C. sorokiniana* needs a variety of nutrients and a certain environment for survival. In addition to the four basic elements; C, N, P, and S, ionic components such as sodium, potassium, iron, magnesium, calcium, and trace elements must also be provided for algal growth.

The environmental conditions affect algae growth and nutrient uptake. Conditions include extracellular pH, temperature, salinity, light intensity.

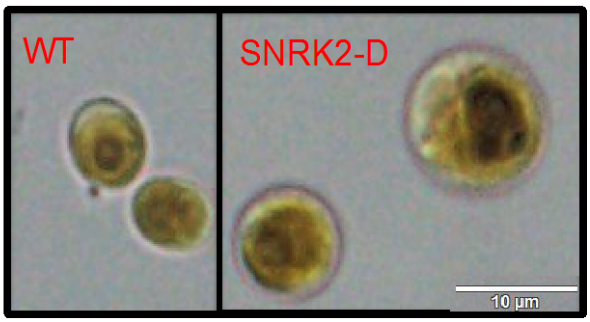
D.2.b. Laboratory studies comparing survival of the subject microorganism and the unmodified parental recipient strain in soil or water samples taken from the release site.

At the time of this submission we have not yet tested the recipient and subject microorganism together in soil or water samples from the release site. We will perform those laboratory tests in the months preceding the commencement of the field trial if approved. We will be able to receive the subject microorganism in-house under our NIH R&D exemption (40 CFR parts 725.234 and 725.235) and perform the water and soil survivability tests in the AzCATI Laboratory. However, and indoor greenhouse trial was performed at the LANL New Mexico site and that data is presented below. The experiment was performed in a controlled environment greenhouse under with all protocols for operation extensively reviewed and approved by LANL Environmental Health and Safety with appropriate secondary containment and monitoring and only specific personnel allowed to interact with the ponds and samples and conduct the experiment. The controlled environment greenhouse is set up for the running of transgenic plants and algae and the appropriate controls to limit any unintentional environmental release.

Laboratory shake flask growth comparisons between recipient and subject microorganism: As shown in Figure 26, the SNRK-2 D line showed higher growth compared to the recipient (wild type) microorganism in both low light and high light shake flask conditions. SNRK-2 D lines exhibited larger cells and demonstrated higher starch content when stained with lugol's staining solution. SNRK 2-D also showed higher photosynthetic efficiency, reduced chlorophyll content and smaller antenna size (high chl a/b ratios) compared to recipient microorganism.

Indoor Greenhouse 50L Minipond Experiment: Based on the performance of the SNRK2 comparative growth analysis for SNRK2-D with WT indoors in flask experiments, an experiment was designed and executed in the LANL's controlled environment greenhouse using small mini-pond reactors at a volume of 50L in June/July 2018. The cultivation experiment was to compare the growth of Cs1412 WT (recipient) and SNRK2 D (subject) in triplicate mini-pond in the greenhouse. The culture volume for each pond was 50 L and the cultures were grown in HS media (pH 8) containing ammonium chloride as the nitrogen source. Sodium carbonate was used to adjust the pH of cultures. An illustration of the experimental plan for the cultivation of Cs1412 and SNRK2 in mini-ponds is shown in Figure 27. Starting inoculum for WT and SNRK2 ponds were grown in spin flasks (culture volume 1.5 L/flask) and two starter mini-ponds (one WT pond and one SNRK2 pond) were inoculated at an OD_{750} of ~0.25. After two days of growth, each starter pond was split into three replicates to create triplicate ponds for WT (average OD_{750}=0.39) and SNRK2 (average OD_{750}= 0.38).

C



Lugol's Staining (◊Day 4, low light)

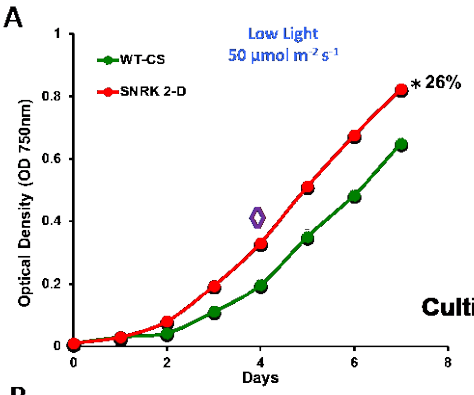


Figure 26: Photoheterotrophic growth of SNRK-2 D transgenic line in HS media (A) Low light intensity. **(B)** High light intensity **(C)** 4 days old WT and SNRK-2 transgenic cells were stained with Lugol's solution for starch accumulation and microscopic imaging. Growth was measured at 750nm. These experiments were performed in triplicates. * indicates significant increase in growth.

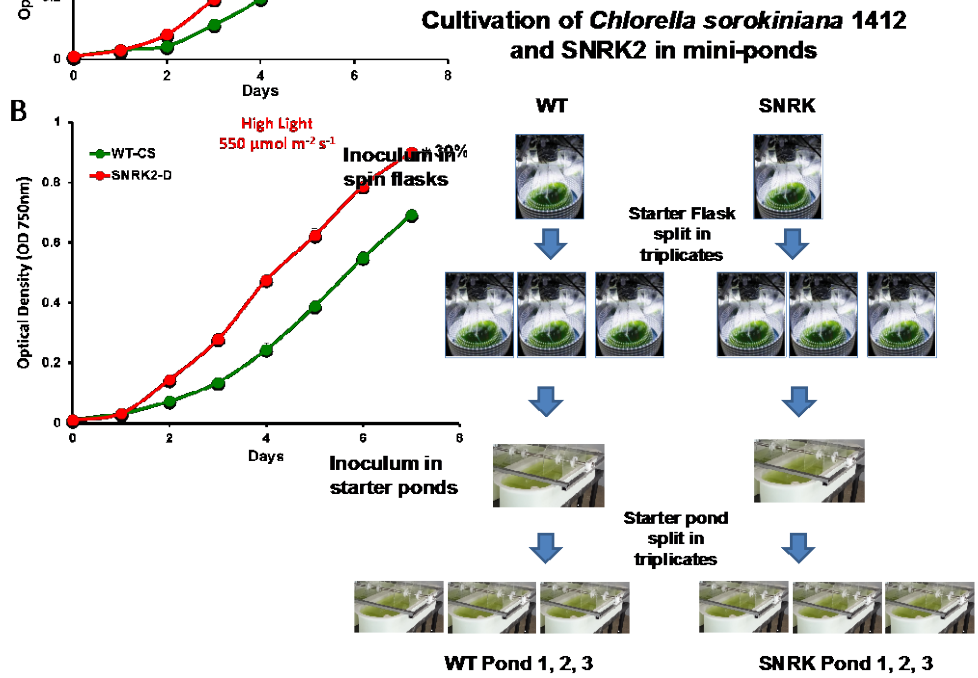


Figure 27: An illustration of the experimental plan for the cultivation of Cs1412 and SNRK2 in mini-ponds in the LANL controlled environmental greenhouse (note that these are not the actual pictures of the spin flasks/ponds from this experiment).

Throughout the experiment samples for optical density measurements, microscopy and biochemical compositional analysis (~2 L) were collected daily from each pond until the cultures reached the stationary phase (nitrogen depletion) and every other day thereafter. Prior to sampling, evaporation loss was replaced with water and after sampling, the sample volume was replaced with nitrogen-free media because of the relatively large samples collected. Nitrogen depletion in the cultures was monitored with ammonia test strips. The growth data for the triplicate Cs 1412 WT and SNRK2 ponds are shown in Figure 28. These results show variability in growth patterns of replicates, which was seen in both WT and SNRK 2D pond replicates. This was first time CS1412 was cultivated in mini-ponds in the LANL greenhouse and the differences observed are being investigated to better understand the differences in the growth patterns of replicates. However, initial results indicate that SNRK 2D phenotype for cell size and starch accumulation is consistent with our previous experiments performed in shake flasks (see Figure 26 and Figure 29). As shown in Figure 29 on day 3 SNRK2 D shows higher carbohydrate accumulation compared to WT. **SNRK-2 D lines showed 21% increase in total carbohydrate accumulation compared to WT on day 3.**

PCR confirmation of SNRK 2 gene stability: From the starting of the mini pond growth experiment we are keeping track of stability of the SNRK gene insertion in SNRK 2-D line. We performed the PCRs for the validation of the gene presence from the starting plate and spin flasks (Figure 30). We have also confirmed the SNRK 2 gene presence in the Day 3 samples. As shown in Figure 28 we have selected days throughout

the experiment for PCR verification to determine the presence of SNRK 2 gene in the SNRK 2D lines and also to keep track of any cross contamination of SNRK gene in the WT ponds.

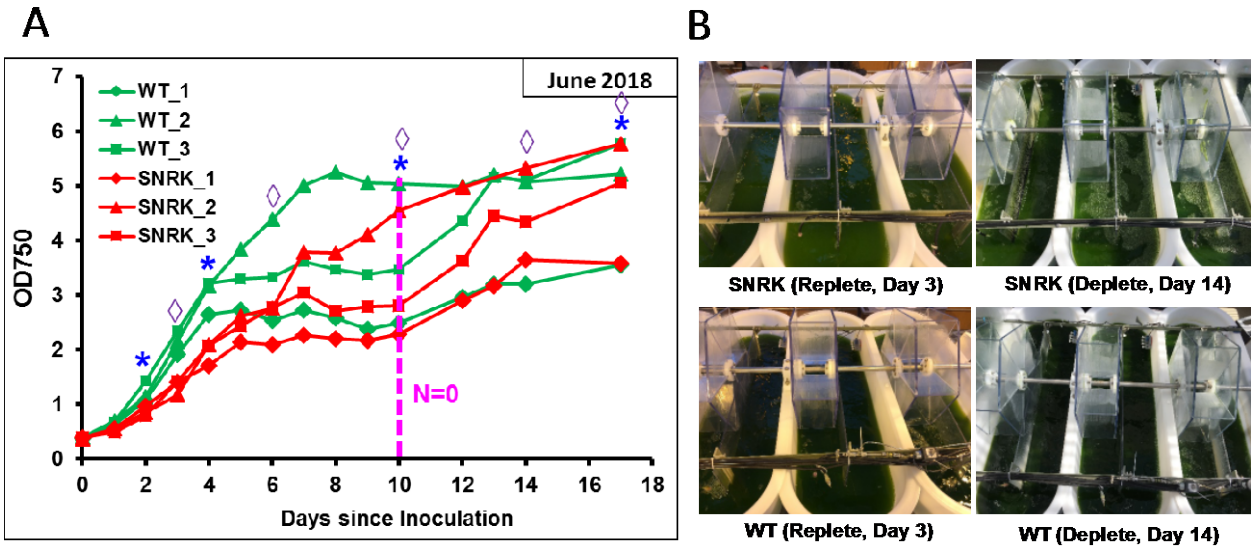
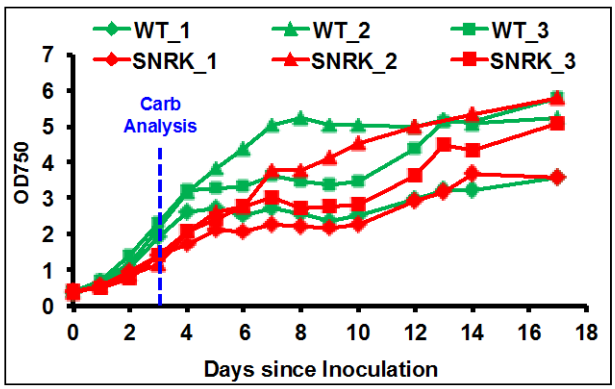
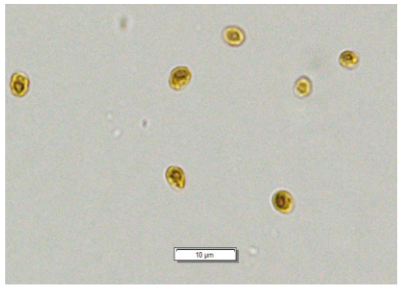
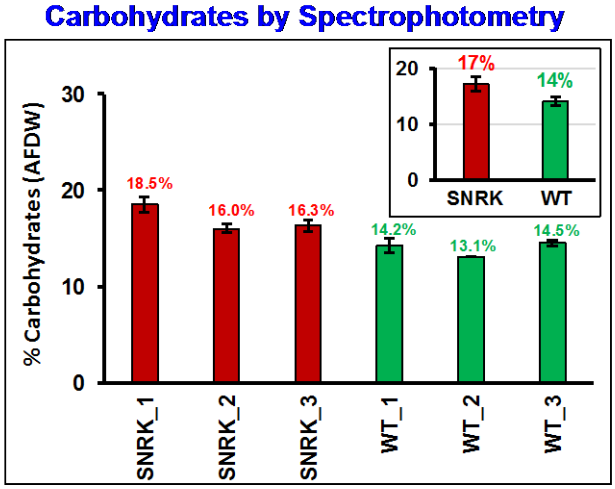


Figure 28: A) Growth data for Cs1412 and SNRK2 ponds. * and ◊ represent the days selected for PCR screening and microscopy. B) Pictures of the Cs1412 and SNRK2-D ponds on day 3 (replete) and day 14 (deplete).







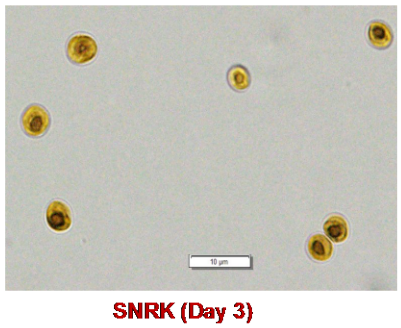


Figure 29: Spectrophotometrically determined carbohydrate contents for SNRK2 and WT biomass collected on Day 3 (N-replete). These data are reported on ash-free dry weight basis.

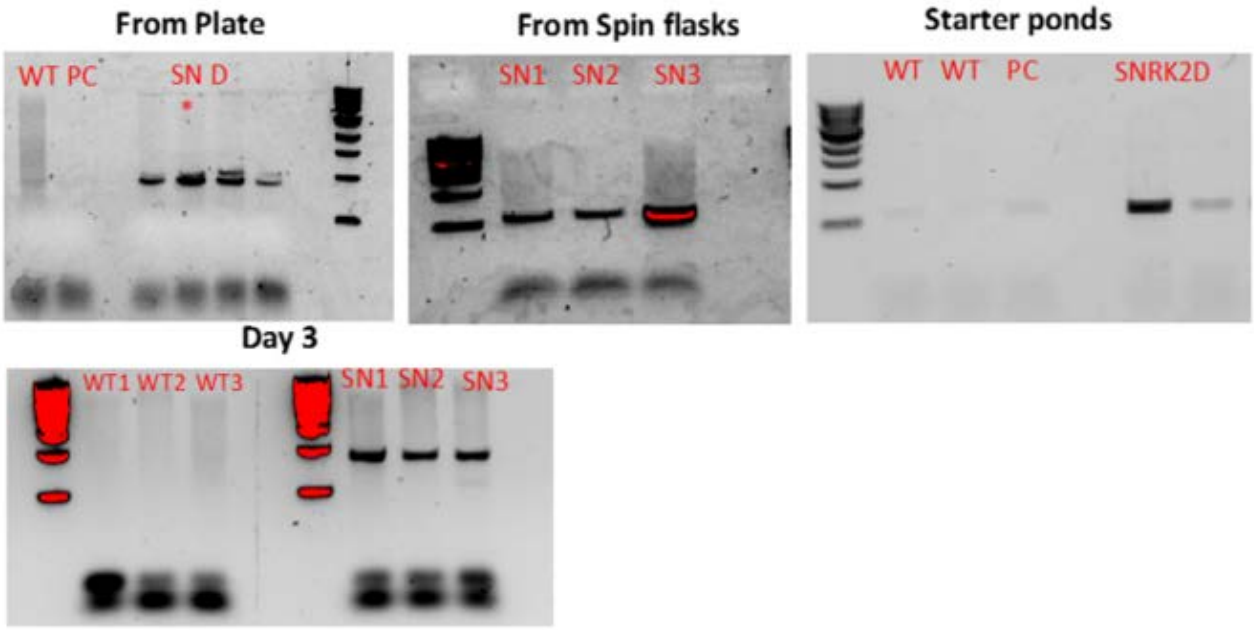


Figure 30: PCR confirmation for presence of SNRK2 in mini-ponds and absence in control ponds.

Catch plate confirmation of algae dispersion: We performed catch plate analysis to check whether algae from the ponds were dispersed in the atmosphere. The catch plates were placed in close proximity to the ponds and along the airflow direction present in the greenhouse. Agar plates with HS media were kept open during the time of sampling for 1-2 hours and after sampling, plates were wrapped and put under light to incubate. If algal growth was present on the plates this would indicate that algae was being disperced into the atmosphere and transported across the greenhouse. As shown in Figure 31, even after one month of operation in the controlled environment greenhouse, there was no algal growth on any of the catch plates, however there was minor fungal growth on the plates.

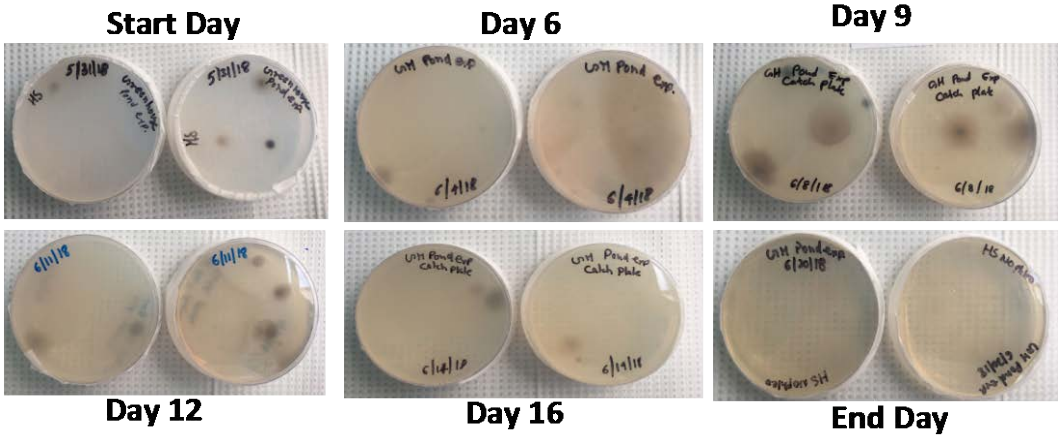


Figure 31. Catch plates from greenhouse showing only fungal growth on the HS media.

D.2.c. Survival/persistence in other environmental media aside from that found at the release site (e.g. other types of soil, water, and/or air).

At the time of this submission we have not yet tested the recipient and subject microorganisms together in water sources found near the AzCATI release site. As described in Section F.1, and shown in Figure 8, the only surface water in close proximity to the site is a water canal that is a source of drinking and potable water. This water eventually makes it way to water treatment facilities to be processed into drinking/potable water. Samples from this source will be collected and we will perform survival/persistence studies with the subject and recipient microorganisms. Historically, this source of water has been a source of algae prospected by ASU, but is generally lacking in essential nutrients to support algal growth without supplementation. We will perform laboratory tests on filtered and unfiltered canal water in the months preceding the commencement of the field trial if approved. We will be able to receive the subject microorganism in-house under our NIH R&D exemption (40 CFR parts 725.234 and 725.235) and perform the local water survivability tests in the AzCATI Laboratory.

D.2.d. Known and predicted environmental conditions that may affect survival, multiplication, dissemination.

C. sorokiniana needs sunlight, CO2, nutrients, and moisture to survive and multiply. We have cultivated the recipient strain for the last 2 years at the AzCATI site in a variety of media types and cultivation configurations from closed photobioreators to open ponds. While we do expect the SNRK2 strain to show improved biomass productivity and composition (higher carbohydrate content) relative to the recipient microorganism, we do not expect to see any difference in dispersion or advantage to survive, multiply, or disseminate for the subject microorganism compared to the recipient microorganism in the surrounding ecosystem, and if so, then we do not expect the genetic changes to show any adverse effects in that ecosystem.

D.2.e. Method of detection and detection limits of microorganism in soil and/or water (particularly in the intended environmental media or release site).

Algae can be cultured from soils from surrounding areas and identified on the basis of 16S and 18S rRNA gene sequences as well as the presence of possible transgenes.

D.2.f. Prevalence of gene exchange in natural populations.

Key terms *C. Sorokiniana** gene transfer* genetic exchange* searched on PubMed did not produce results indicating the prevalence of gene exchange of the recipient microorganism in natural populations.
Prior to initiating the field trial experiments we will sample soil from the field test site to determine what organisms may be present. We have been cultivating the recipient microorganism for over 2 years but have cultivated other *Chlorella* sp. for over 11 years, in particular in close proximity to the location of the miniponds to be used for this experiment. We intend to utilize a soil DNA sampling kit (e.g., PowerLyzer PowerSoil DNA Isolation Kit, Mo Bio Laboratories), extract DNA from soil samples across the field site and use the extracted DNA in one of two PCR reactions. In the first reaction, we will use primers specific to conserved regions within the 18S ribosomal DNA region including sequences found in most eukaryotes. In a second reaction, we will use primers specific to conserved regions in the 16S ribosomal DNA sequence found in most prokaryotes. We have been monitoring pond ecology in a similar fashion for the past 6 years as part of a separate DOE funded program (ATP^3) and will use that as a reference baseline, extending this to understanding the microbial community within the soil. We will perform a similar analysis on the local surface water from the canal system to the east of the Polytechnic campus.

D.3 Summary

As described in Section *C* and *D, Chlorella* sp. in general are quite ubiquitous in the environment and an important primary producer and food source for higher trophic levels. There is no evidence in the literature of the recipient microorganism producing toxins, antibiotics, xenobiotic or degradative gene products. A homology search of the subject microorganism's intergeneric genes against known allergens did not return any significant results with no identity matches greater than 35%. While we expect some ability for the recipient and thus the subject microorganism to be capable of dispersion through airborne transfer, there is little information on this for this specific strain and thus performing these field trials as proposed is an important next step with relatively low risk.

E. PREDICTED PRODUCTION VOLUME, BYPRODUCTS, USE AND CONSUMER EXPOSURE

E.1 Information on Production Volume

E.1.a. For batch processes, the batch volume, the maximum cells/batch, and the maximum CFU/batch that are likely to be produced in the first year of production; and the maximum cells/batch and the maximum CFU/batch that are likely to be produced during any 12-month period during the first three years of production. Provide also the maximum cell density of the fermentation broth in CFU/ml, the number of hours needed to produce each batch, and the number of batches per year.

E.1.b. For continuous processes, the maximum cells/day and CFU/day that are likely to be produced during any 12- month period during the first three years of production, the # hrs/day and # days/yr, and the maximum cell density of the fermentation broth in CFU/ml.

The purpose of this trial is to compare the subject and recipient microorganisms when cultivated in outdoor miniponds with respect to their ability to produce biomass, and also to identify any increase or decrease in biomass productivity under biotic and abiotic stress from the environment. In addition, we will evaluate the ability of the subject and recipient microorganisms to disperse and migrate into trap ponds at the test location site as well assess the potential risk to the local ecosystem by evaluating the ability for the subject and recipient organism to grow in local ecosystems gathered and placed in those trap ponds. The trial will take place at the Arizona Center for Algae Technology and Innovation (AzCATI) testbed facility located at Arizona State University's (ASU) Polytechnic campus in Mesa, Arizona and will last approximately 60 days. ASU's Polytechnic campus is approximately 40 km southeast of Phoenix and the AzCATI testbed and laboratory are located in the southeast corner of the campus (see Figure in section *F.1*). The ideal start date for this experiment is on or about September 1, 2017. We will utilize miniponds already in use at the AzCATI testbed. We will utilize 6 ponds total. These are 800-1000L (depending on operating depth of the water) and approximately 4.2 m^2 in surface area for each pond (see Figure 7, panel C, in Section *F.1*). They are on raised stands and will be placed within secondary containment constructed of a wooden frame acting as a berm and overlaid with industrial liner appropriate for this application. The volume capacity of the secondary containment will be sufficient to hold 1.5X the volume of the ponds under operation during the experiment (e.g., 7.2-9.0 m^3 total). Note that these ponds have been operated at AzCATI for 5 years and we have never experienced a catastrophic failure leading to uncontrolled release of culture. Ponds will be inoculated at a starting biomass density of 0.1-0.2 g/L on an ash free dry eight (AFDW) basis. Ponds will be harvested when the density reaches 0.4-1.0 g/L (AFDW) with approximately 50-80% removal of the total biomass present. Ponds will then be brought back to the starting volume/depth, replenished with additional nutrients/media components and allowed to grow out again to the target harvest density. Biomass samples during the run will be isolated as a slurry through dewatering with a centrifuge. Small volume samples (1-4 L) will be centrifuged in the laboratory. Larger volumes (>500L) will use a small continuous centrifuge immediately adjacent to the

pond pad and will also be performed under secondary containment protocols and appropriate PPE. Operating in this semi-continuous fashion, we would expect to repeat this cycle 1-10 times reaching first harvest in 5-15 days.

E.2 Information on Byproducts

In addition to information on the amounts of viable cells produced per year, it may also be helpful to know about the concentration of proteins, DNA, or other materials produced as byproducts of the manufacturing process, if such material may pose hazards to humans or the environment. Please indicate amounts of byproducts which may pose hazards to man or the environment as was done for viable cells under item E.1.

There will be no viable byproducts as a result of this experiment. All biomass, whether used to generate cultivation data (dry weight, photosynthetic efficiency measurements, OD750/680, nitrogen/phosphorus analysis, etc.), proximate composition, and hydrothermal liquefaction (HTL) bio-oil will be inactivated as part of the analysis process and disposed of properly to ensure that the microorganism can no longer replicate. As referenced in previous sections, it is not expected that the DNA or proteins contained within the subject microorganism will cause harm to humans or the environment.

E.3 Use Information and Consumer Exposure

E.3.a. Describe the intended use(s) of the microorganism for the particular processes for which it is intended (such as waste degradation) or products it is intended to produce (such as enzymes for detergent use

The subject microorganism will be used in this experiment for two purposes: 1) to investigate the translatability of phenotypes from lab to field and 2) to understand how microalgae migrate and affect natural plankton communities.

E.3.b. Estimate the percent of the production volume for each use.

Ponds will operate between 800-1000 liters and between 20-80% of the biomass will be harvested periodically and dewatered to a slurry and freeze dried or stored as frozen. We estimate that <5% of the biomass will be collected for all other types of data analysis (e.g. dry weight, photosynthetic efficiency measurements, OD750/680, nitrogen/phosphorus analysis, proximate composition and molecular fingerprinting, etc.). At the conclusion of the experiment, all of the remaining biomass will be harvested and disposed of as described in Section G4.d.

E.3.c. Estimate the concentration of the microorganism in the formulations for each use identified.

E.3.d. Generic use if the use is claimed as CBI.

E.3.e. Based on the intended use(s), identify products in which the subject microorganism or its product will be present.

E.3.f. Presence of subject microorganism in consumer products (estimate amount if possible).

E3.c-f. The purpose of this TERA submission is to perform a time limited, defined scope R&D experiment; therefore we will not be creating any formulation, product, commercial or otherwise, or have any other uses for this microorganism. We will use the data obtained from this experiment to inform future R&D activities and future TERA/MCAN submissions.

F. PREDICTED RELEASES DUE TO MANUFACTURING OF THE SUBJECT MICROORGANISM AND WORKER AND CONSUMER EXPOSURES TO THE SUBJECT MIRCROORGANISM

F.1 Industrial Sites Controlled by the Submitter

F.1.a. Operation description

The activities described in this TERA will take place at ASU's AzCATI testbed facility, located at 7418 Innovation Way South, Mesa, Arizona 85212. The AzCATI site is located across the street from the ISTB-3 laboratory building on the Polytechnic Campus of ASU. It occupies approximately 4.0 acres of which 0.6 acres contains production facilities with a cultivation capacity of 300,000 L. The site includes 3 greenhouses and 2 head-houses and is under the direct control and supervision of AzCATI personnel. The site is fenced-in with restricted access limited to authorized personnel. The site contains 32 mini ponds (0.5-5 m^2, and volume of 0.1 – 1.5 m^3), 2 medium raceways (60 m^2, volume up to 15 m^3), an ARID™ raceway (400 m^2, volume up to 40 m^3) and 1 large raceway (500 m^2, volume up to 125 m^3). In addition, the site has 50+ closed photobioreactors in different configurations (flat panel, hanging bag, plastic/glass tubular) in volumes from 0.025 – 1.5 m^3). For the activities described herein, we will also utilize at least 6 and up to 12 trap ponds for the dispersion testing. These will contain either water from a surface source described in section *G.2.d* or synthetic media matching what is used on the proposed cultivation trials as described in section *G3.* There is also downstream processing equipment for dewatering algal biomass from the ponds including sedimentation, dissolved air floatation, centrifugation and membrane filtration up to a pilot scale.

F.1.a.1. Identity - identity of the site at which the operation will occur to include the name, site address and city, county, state and zip code.

The field experiment will take place at the AzCATI Testbed facility, which is located on the Polytechnic campus of Arizona State University in Mesa, AZ located at 7418 Innovation Way South, Mesa, Arizona 85212. Its proximity to Phoenix, AZ is shown in Figure 32.

F.1.a.2. An indication of whether the operation is best described as manufacturing, processing, or use

The operation described herein is best described as experimental use.

F.1.a.3. Process description

The recipient microorganism will arrive to AzCATI laboratory under the PACE Material Transfer Agreement on agar plates and will be scaled up through small shaker flasks, to 800 ml bubble columns to 2'x2' flat panel photobioreactors with a 2.0" light path operated at max volume of 10L. These activities are subject to ASU's NIH R&D exemption (40 CFR parts 725.234 and 725.235). The AzCATI team has over 9 years of experience running cultivation trials from 50 -2000 L scale for GM algae and cyanobacteria in closed systems both indoors and outdoors and have all necessary protocols already in place and approved to meet the R&D exemption for the scale up in closed systems. Once we are at a density of 2.5 g/L in the 10L flat panel reactors for the subject and recipient microorganisms we will combine multiple panels to have sufficient biomass of approximately 240-300 g on an AFDW basis (e.g., ~12-14 panels for each subject and recipient microorganism) in a total volume of 120-140 L for each strain. We will combine culture into approved containers with secondary containment to transport seed culture from the laboratory across to the field site (approximately 100 m). Triplicate ponds will be inoculated to a starting density of at least 0.1 g/L at a depth of 800-1000 L. Sampling of the ponds will be as described in Section *G.4.d)*. Biomass not used for sampling or downstream conversion will be discarded as described in Section *G.4.d.* All ponds and equipment used in operation will be cleaned with a 5% bleach solution. At no point in the

process/experiment will there be an intentional release of the subject microorganism into the environment. Mitigation strategies to minimize exposure to and accidental release of the subject organism are described in more detail in Section *G.1.b* and expected exposure levels are shown below in Table 7.

Table 7: Expected exposure of workers to subject microorganism.

Process Step	Personnel Protective Equipment (PPE)	# of workers	Maximum duration (hrs/day)	Maximum duration (days/year)
Scale up and initial inoculation	Proper PPE*	2	2	30
Routine Pond Monitoring	Proper PPE*	2	2	60
Grab Samples (scope, OD, AFDW, qPCR, proximate analysis, etc.)	Proper PPE*	2	2	60
Sample Processing	Proper PPE*	3	6	45
Pond harvests/resets	Proper PPE*	2	6	10
Terminating Experiment	Proper PPE*	2	4	3

F.1.b. Results of any personnel or area monitoring conducted during the production process.

The field site is secure with controlled access and has CCD TV monitoring. However, there is no ability to monitor personnel other than our normal attention to our student and staff's general health and welfare.

F.1.c. Environmental Release and Disposal

There is no intended release of the subject microorganism in this experiment. All workers involved in the field experiment will follow established labeling and disposal protocols to ensure that all subject microorganism samples are handled appropriately. As described in Section G4.d, samples that have been collected from the site will be inactivated by treatment with a 5% bleach solution and autoclaved. Bulk cultures will also be inactivated by treatment with bleach and disposed of into the sites sanitary sewer per existing protocols for large scale culture disposal.

F.2 Industrial Sites Controlled by Others

At this time, the subject microorganism will only be released at a site completely controlled by the applicant and will not be distributed to any industrial site controlled by any unrelated party.

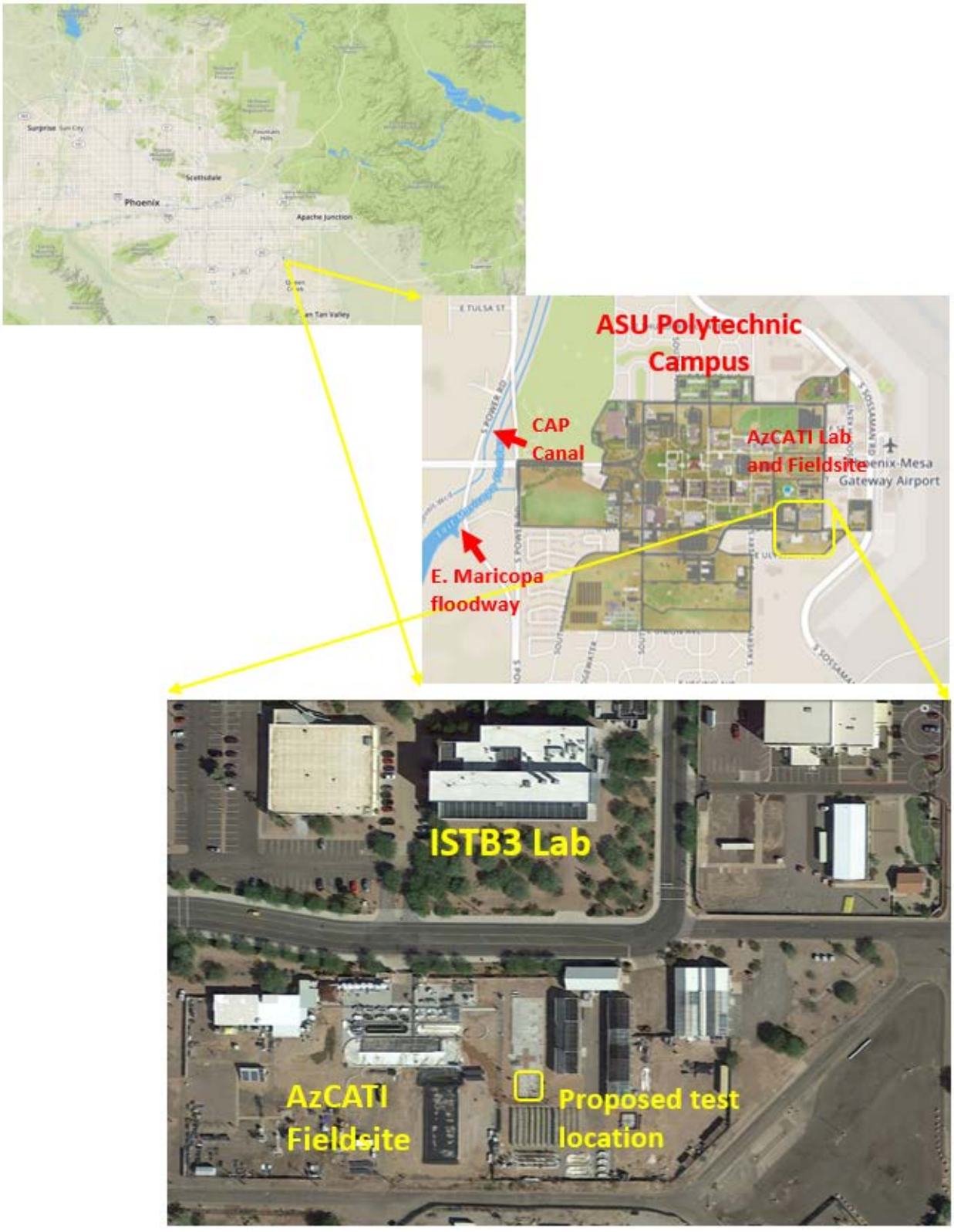


Figure 32: ASU's Polytechnic campus is located in Mesa, AZ approximately 40 km southeast of Phoenix, AZ in the southeast valley of the Phoenix metropolitan area. The AzCATI laboratory and field site are located in the southeast corner of the Polytechnic campus.

G. INFORMATION APPLICABLE TO THE FIELD TESTS FOR THE SUBJECT MICROORGANISM

G.1 Objectives

G.1.a. List the objectives of the field trial and describe the rationale which requires the environmental release of the subject microorganism.

Bioenergy from algae has the potential to contribute substantially to the nation's renewable energy future, but significant challenges surround the transition to a commercial scale algae farm (Quinn, et al. 2015 and Laurens, et al. 2017). Some of the barriers can be attributed to disparate literature reports on productivity and compositional estimates for algae production. In addition, the lack of agronomic data on algae cultivation makes future-year projections difficult and unreliable. Increased research and development as well as near- or at-scale demonstrations are needed to improve algal productivity, pond performance, reduce risk and uncertainty in deployment of an algal biofuels process, and address and validate the data gap between assumed and actual experimental value (EERE-BETO MYPP (2016) and ANL; NREL; PNNL (June 2012)). This is as much, and even more so the case for GM approaches seeking to improve productivity and enable valuable co-products, as there is added uncertainty and risk to responsible deployment of GM algae, as there was for deployment of other GM crops. The eventual large scale deployment of GM algae by its nature implies there will be potential for, and in fact the eventuality of release into the environment. This requires a deliberative, responsible and stepwise approach to assessing those risks. And that requires scientific data from the field, on not only the assessment of whether those traits engineered in the laboratory will translate when deployed to the field, but careful and controlled experiments to determine the fate of GM algae in the environment through open release trials. The experiment proposed here has two main objectives and builds on the only previous open release GM algae trial conducted to date (UCSD/Sapphire) which provided the framework for this subsequent TERA application:

1) Evaluate the ability to translate laboratory phenotypes of GM algae from indoors to relevant outdoor conditions in small, commercial style high throughput raceway ponds.
2) Characterize the potential risks and ecological impact (dispersion, invasion) of those GM microalgae.

G.1.b. Describe the possible benefits and risks of the proposed field test.

The primary benefit of the field test is an understanding of the ecological impact of GM microalgae on neighboring ecosystems and validation that genetic modification of *C. sorokiniana* can translate to relevant outdoor conditions, enabling more economical deployment and thus, commercialization of algae based biofuels and bioproducts.

As with the previous open release application executed by UCSD/Sapphire in 2013, the primary risk of this experiment is that there can be a loss of containment via a spill. AzCATI and ASU have a long history of conducting large scale testing (from 50-2000L) of GM algae in closed systems performed under ASU's NIH R&D exemption (40 CFR parts 725.234 and 725.235) and the risk mitigation for this TERA we will undertake is similar to previous experiments in several ways: 1) the miniponds to be used have a long history of use (>5 yrs.) without a single instance of a catastrophic structural failure (i.e. leak) and all systems will be tested and thoroughly checked for leaks prior to and during experimental use, 2) the miniponds will be contained within a 9m x 11m perimeter that is underlined by a mesh-reinforced, puncture-resistant, UV-resistant pond liner, 3) all workers will be outfitted with the proper PPE, and 4) all on-site workers will be trained to follow established protocols set in place by ASU/AzCATI and approved by EH&S/Biosafety. These mitigation strategies taken together with our history of successful GM closed

reactor trials will help ensure that loss of containment is a minor risk – but if it does occur, there will be practices in place to minimize environmental exposure. Migration of the subject microorganism into neighboring pond(s) represents another risk for our site. To mitigate this risk, we will sample weekly from other operating ponds at the field site. We will analyze the samples for the subject microorganism via qPCR analysis. Once the subject microorganism has been detected in a neighboring pond, we will survey that pond three times per week. If the detected subject microorganism increases by a factor of 1000 from its original titer (indicating that the subject microorganism has established itself and is actively replicating) in any of the neighboring ponds, we will consult the TSCA office to determine further actions and will be ready to terminate the experiment, if necessary.

G.1.c. Approximate start date and the duration of the field test.

The intended start date is on or about October 15, 2018 and will take place for approximately 60 days. Unforeseen technical/weather issues could cause us to modify the exact start and finish dates.

G.2 Nature of Site

G.2.a. Location and size of the test area

The experiment will be carried out at ASU's AzCATI Testbed facility, which occupies approximately 4.0 acres. The area within AzCATI that we will use to cultivate the subject microorganism is a minipond array of 6 ponds, sitting in a 9m x 11m containment area with trap ponds placed on the AzCATI Testbed site as shown in Figures 33 and 34 in Section *G.3.b*.

G.2.b. Describe why this site was selected (relying on items such as those listed in 2.c. - 2.g. below).

This site was purpose built for performing algae cultivation trials at various scales outdoors. There has been active outdoor cultivation work since 2006, with major expansions in 2008 and again in 2011. It is collocated on the polytechnic campus of Arizona State University and immediately adjacent to a 30 square foot state-of-the-art laboratory facility. We have an 8 year history of working with genetically engineered, transgenic algae and cyanobacteria under closed cultivation conditions, which includes both indoor and outdoor cultivation, including at scales >1000L. This site also serves as the lead site for a Department of Energy funded national testbed network (separate funding than for the PACE consortium).

G.2.c. Describe the history of site use. Describe any prior agricultural uses of the site, if applicable. Include items such as cropping history, tillage management systems, fertilizer and pesticide applications, and other factors, conditions, or practices which might influence characteristics of the site. If the site is not an agricultural site, describe other prior uses of the site that would influence survival, distribution, and effects of the subject microorganism such as wastes present in surface or subsurface soils, surface or underground installations, etc.

ASU's Polytechnic campus used to be part of the Williams Field Airforce base. While the surrounding areas immediately adjacent still are utilized for agricultural purposes, the land the testbed sits on was part of the formal airbase and was used for road and building since the 1930's. The testbed site has been used for outdoor wild-type algae cultivation since 2006. Prior to that it was a vacant lot when ASU acquired the property in 1996. The soil type across the campus and in particular on the field site is of a hard packed caliche type (i.e., soil that is cemented together by calcium carbonate) that is alkaline and is very common in southern Arizona.

G.2.d. Describe physical characteristics of the site related to surface and ground water such as distance to surface or ground water including public and private drinking water sources. For groundwater include

temperature, flow velocity, dissolved oxygen, dissolved organic carbon, suspended solids, direction of flow, volume, depth, width of aquifer, and pH.

In the Phoenix area, water comes from one of three sources: groundwater (either from a private well or one operated by a city or private water company), the Salt and Verde rivers (delivered by Salt River Project- SRP) or the Colorado River (brought here by the Central Arizona Project – CAP) in a canal that stretches 336 miles from Lake Havasu to Tucson. The Polytechnic campus' water source is a municipal source from the City of Mesa. The site is not within close proximity to any natural lakes (two largest reservoirs are Roosevelt Lake (70 km to the northeast) and Lake Pleasant (80 km to the northwest). There is a CAP/SRP water canal that runs along the west side of the campus approximately 1.5 km due west of the field site see (Figure 8). We will utilize this as one of the "wild community" sources of native ecosystems for "natural" bodies of water for our assessment of invasion potential.

G.2.f. Evaluate the possibility of dissemination to adjacent ecosystems and other characteristics of the site that would influence containment or dispersal e.g. relation to flood plain, slope, average wind speed and direction, annual rainfall.

As mentioned above, the site is far removed from any surface waters, except the CAP canal directly to the west. During the time of this experiment, prevailing winds are primarily from the west and blow east/southeast. Rainfall during this time of year averages 0.6-0.7" month (Sept-Nov). In addition, to reduce the risk of inadvertent contact by birds, which could then transport algae great distances from the test site even further for dissemination to neighboring ecosystems, we will employ bird netting over the ponds to mitigate this risk.

G.2.g. Identification and description of nontarget human and nonhuman populations of concern that may be exposed, e.g. distance to nearest dwellings and population density around site.

The. The Polytechnic campus of ASU is very urban and sits at the eastern edge of Gilbert and southeastern part of Mesa. There is an airport directly adjacent to campus to the east and open desert, some farmland to the south/south east. There are residences within 1 mile of the field site to the northwest and west and southwest and are mainly campus housing. There are no sensitive natural ecosystems within 10 mi of the campus. As mentioned in *G.2.f.*, a physical barrier (bird netting) to further mitigate any risk to inadvertent contact and transport by wildlife.

G.3 Field Test Design

G.3.a. Rationale for field test design, description of proposed statistical analyses, and explanation of how the statistical analysis will answer the field test objectives.

The experiment proposed here has two main objectives and builds on the only previous open release GM algae trial conducted to date (UCSD/Sapphire) which provided the framework for this subsequent TERA application:

1) Evaluate the ability to translate laboratory phenotypes of GM algae from indoors to relevant outdoor conditions in small, commercial style high throughput raceway ponds.
2) Characterize the potential risks and ecological impact (dispersion, invasion) of those GM microalgae.

To achieve the first objective, subject and recipient microorganisms will be scaled up from agar plates or slants, taken through 100 ml then 300 ml shaker flasks then into 800 mL bubble columns. Columns are run under continuous cool white fluorescent light (~ 150 µmol m^{-2} s^{-1}) with continuous bubbling with 2% CO2/air (v/v). Once a culture density of approximately 0.8 g L-1 ash free dry weight (AFDW) is reached in

four columns, they will be split four ways to a total of 16 columns. Once the column seed cultures reached a density of approximately 2.5 g/L AFDW, 14 of the columns will be used to inoculate fourteen 2' x 2' vertical flat panel reactors each with a 2" light path and a nominal 10 L volume. Two of the columns will be retained as backup seed culture in the event they are needed for pond reset due to a failure (i.e., culture crash). The same lighting and 2% CO2/air v/v will be used for the panels. PH will be ~ 7.0-7.2 under the indoor conditions and also will be the target pH for the outdoor miniponds. The panels will be started at an initial target density of approximately 0.15 g/L AFDW. Once the panel cultures reach a minimum density of 2.5 g/L, they will then be combined into a single master batch inoculum and used to inoculate three outdoor ponds for each subject and recipient microorganism respectively, at a minimum target density of 0.1 g/L AFDW. Typical duration for seed culture scale-up from columns through to panels and finally to outdoor ponds will be ~28 days. An example of typical growth for the recipient microorganism under the scale up process and outdoor pond operation conditions to be employed are shown in Figure 7 in Section *F.1.a.3*. Typical media recipes for the DOE1412 strain are shown in Table 8 and will be the basis for the outdoor experimentation (likely HS or BG-11 variants).

Table 8: Media recipes used for DOE1412 cultivation at AzCATI. Target media composition for the field trail will be based on the HS or BG-11 recipes.

	HS Media Background (From LANL sheet)		
Media Component	g/L media final	mM	
H_3BO_3	0.00125	2.02E-02	
$ZnSO_4 \cdot 7H_2O$	0.00125	4.35E-03	
$MnSO_4 \cdot H_2O$	0.00038	2.25E-03	
$CoCl_2 \cdot 6H_2O$	0.00025	1.05E-03	
$Na_2MoO_4 \cdot 2H_2O$	0.00025	1.03E-03	
$CuSO_4 \cdot 5H_2O$	0.00008	3.20E-04	
$MgSO_4 \cdot 7H_2O$	0.1	4.06E-01	
$CaCl_2 \cdot 2H_2O$	0.05	3.40E-01	
*$FeCl_3 \cdot 6H_2O$	0.1	3.70E-01	
*Na citrate	5	2.34E+01	
NH4Cl	0.25	4.67	N:P
$NaH_2PO_4 \cdot 2H_2O$	0.72	4.62	0.7
*K_2HPO_4	0.36	2.07	

	BG-11 Media Background (ATP3)		
Media Component	g/L media final	mM	
H_3BO_3	0.00286	4.63E-02	
$ZnSO_4 \cdot 7H_2O$	0.00022	7.65E-04	
$MnCl_2 \cdot 4H_2O$	0.00181	9.15E-03	
$Co(NO_3)_2 \cdot 6H_2O$	0.000049375	1.70E-04	
$NaMoO_4 \cdot 2H_2O$	0.00039	1.78E-03	
$CuSO_4 \cdot 5H_2O$	0.000079	3.16E-04	
$MgSO_4 * 7H_2O$	0.075	3.04E-01	
$CaCl_2 \cdot 2H_2O$	0.036	2.45E-01	
$C_6H_8FeNO_7$	0.006	2.28E-02	
Na_2CO_3	0.02	1.89E-01	
$C_6H_8O_7$	0.006	3.12E-02	
$C_{10}H_{16}N_2O_8$	0.001	3.42E-03	
NH4Cl	0.267	5.0	N:P
K_2HPO_4	0.0543	0.31	16.0

	F/2 Trace (ATP3)		
Media Component	g/L media final	mM	
$ZnSO_4 \cdot 7H_2O$	0.000022	2.20E-02	
$MnCl_2 \cdot 4H_2O$	0.00018	1.80E-01	
$FeCl_3 \cdot 6H_2O$	0.00315	3.15E+00	
$CoCl_2 \cdot 6H_2O$	0.00001	1.00E-02	
$NaMoO_4 \cdot 2H_2O$	0.0000063	6.30E-03	
$CuSO_4 \cdot 5H_2O$	0.0000098	9.80E-03	
$C_{10}H_{16}N_2O_8$	0.00436	4.36E+00	
Oceanic/Crystal Sea salt	35	N/A	
NH4Cl	0.26747	5.0	N:P
NaH_2PO4	0.0375	0.31	16.0

We will evaluate the biomass accumulation differences at each stage and compare the growth rates using statistical software (e.g., JMP Pro 13.0.0, *SAS Institute USA)* with biological replicates (minimum N = 3) and each measurement run in triplicate. AzCATI has a long history of conducting field trails and comparing one strain against another and that same overall approach to statistical rigor will be used here. The types of sampling we will do is listed in Table 9 in Section *G.3.c.*

Throughout scale up and the outdoor trial, we will confirm at each transfer stage we have the correct subject and recipient microorganism through PCR with specific primers developed allowing discrimination from other *Chlorella* sp., including the recipient, as well as of course other algal species. GM algae are detectable *in situ* by unique intrinsic DNA sequences/and or inserted genetic markers (Henley et al. 2013). For our second objective, characterizing the dispersal and invasiveness of the subject microorganism in the local ecosystem, we will employ water traps external of the miniponds as means to assess the dispersal capability by gathering algae samples that escape the mass cultivation and concentrate these algal cells through centrifugation or filtration of the media/water sample. PCR amplification of the DNA extracted from the concentrated cell sample can indicate the presence or absence of the GM algae using the PCR amplification procedure outlined in Liu et al. 2013 with primers specific to the inserted transgenes. Prior to mass cultivation, a qPCR experiment will establish the minimum detection limit of the marker of interest (see Table 3) for the subject microorganisms to establish the minimum number of organisms required to see a positive result in the screening. With this calibration, we will be able to monitor the trap ponds as well as other cultivation systems that will be

running on the field site and we'll be able to back calculate the number of cells/ml. Minponds, trap ponds and other AzCATI field site ponds will be sampled at the rates shown in Table 9.

G.3.b. Submit a diagram of the plot layout and describe the procedure for randomization of the test plots.

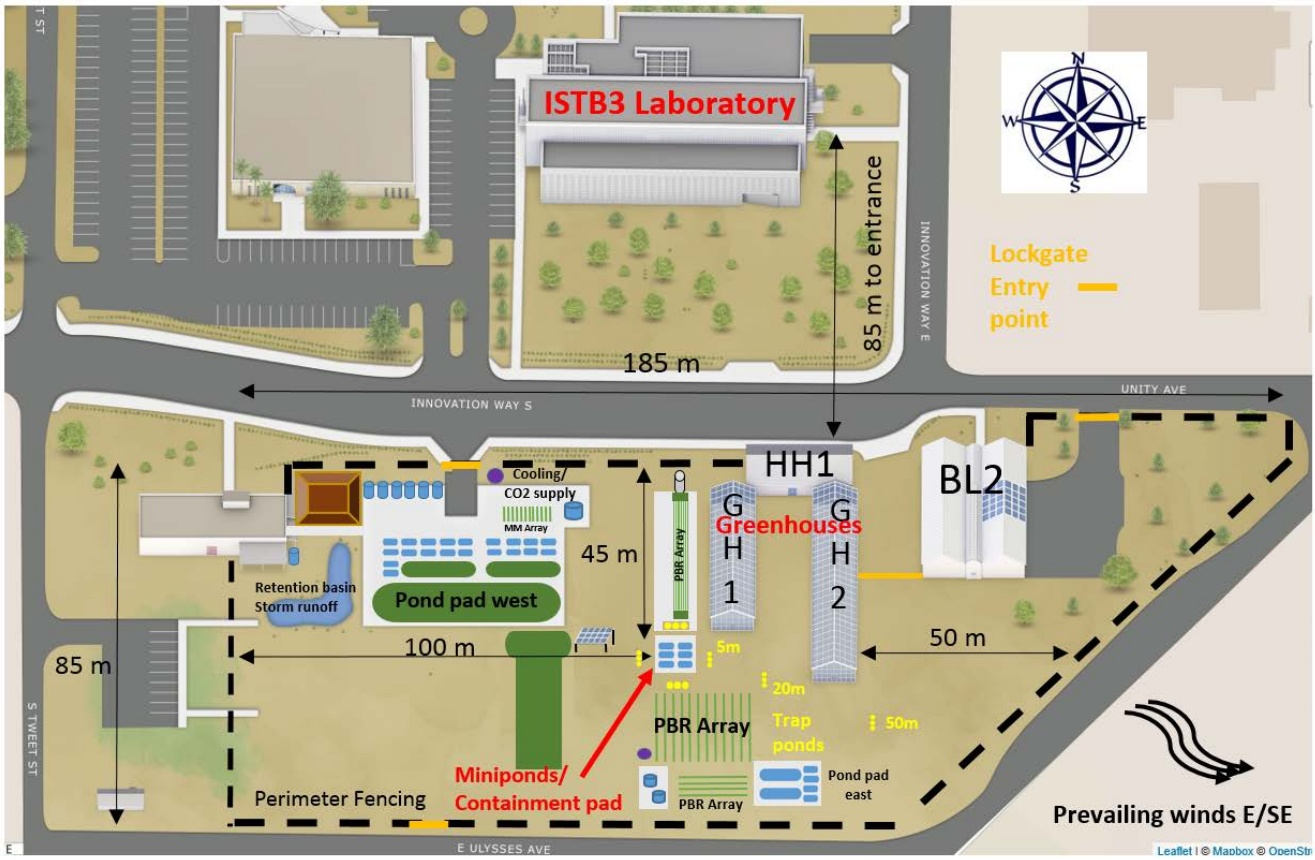


Figure 33: ASU's AZCATI testbed schematic showing general dimensions to perimeter fencing and to the ISTB3 laboratory, location of containment ponds relative to other cultivation systems on site and proposed location of trap ponds placed along path of expected prevailing winds during proposed time period of the experiment.

Figure 34: AzCATI GM containment pad for GM algae testing. Left picture is leak testing the pad, middle picture shows disposal/harvest tank area, and right panel shows ponds placed within the pad. Bird netting will be added either over individual ponds (likely) or over entire area.

G.3.c. In a table: summarize the types of samples that will be collected, where the samples will be taken, how frequently each type of sample will be taken, and how it will be stored prior to analysis

Table 9: Measurement type, sample frequency, source and storage for the experiment.

Measurment	Grab Sample Amount (if applicable)	Frequency of Sampling	Place of Sampling	Storage Conditions
Pond pH/Temperature	N/A	daily	containment miniponds	N/A
$OD_{750/680}$	100 ml (same grab sample as AFDW)	daily	containment miniponds	N/A
DW/AFDW	100 ml (same grab sample as OD)	daily	containment miniponds	N/A
Nutrient Levels in Media (N:P)	5 mL (Supernatunt from OD/AFDW)	daily	containment miniponds	N/A
Microscopic Observation (culture health)	5 ml (same grab sample as OD/AFDW)	daily	containment miniponds	N/A
Pulse Amplitude Modulation (PAM: Photosynthetic efficiency)	10 ml (liquid)	3x/week	containment miniponds	N/A
Pond samples (qPCR)	10 ml (pellet)	daily	containment miniponds	-20°C
Proximate Analysis (total FAME, Total Protein, Total Carbohydrate, ASH)	1-4 L	1-3x/week	containment miniponds	-20°C (freeze dried)
In-situ YSI 5200 sensors (pH, pond water temperature, salinity, % oxygen saturation)	N/A	15 minutes	containment miniponds	N/A
Environemntal (RH, Air Temp, wind speed, wind direction, total irradiance (W/m^2), PAR	N/A	Hourly	weather station	N/A
Precipitation	N/A	Hourly	PHX/Gateway Airport	N/A
Trap pond samples (qPCR)	10 ml (pellet)	3x/week	Trap Ponds	-20°C
AzCATI pond/pbr samples (qPCR)	10 ml (pellet)	weekly	AzCATI ponds/PBRs	-20°C
Bulk harvest (HTL)	500-800 L	1x every 1-2 weeks	containment miniponds	-20°C (freeze dried)

G.3.d. Describe the proposed management of the site. If the subject microorganism is intended for an agricultural application, describe the planting and spacing of the test crop, width and placement of border rows, and, pesticide applications, crop rotation, crop harvesting schedule, etc.

The field site and laboratories to be used are managed by AzCATI as part of ASU. Dr. John McGowen, Principal Investigator for ASU under the PACE consortium, is the Director of Operations for AzCATI and responsible for the day to day operations and will oversee all aspects of the activities proposed under this TERA. Access to the site is restricted to AzCATI personnel and those approved by ASU. The outdoor site is staffed by 1 full-time supervisor, 2 full-time technicians and multiple part-time undergrad students.

G.4 On-Site Containment Practices

G.4.a. Describe the procedures that will be followed for packaging and transporting the subject microorganism to the site.

Initial receipt of the subject microorganism will be on slants/plates and arrive sealed in secondary containment and opened only in the laboratory to begin characterization and scale-up. All samples collected or generated that need to go between laboratory and field site will be labeled with designated

information including, batchID, source location, date, strain ID, collectors name and purpose of the sample. Any sample transported between laboratory and the field site locations will be in sealed container and placed within secondary containment.

G.4.b. Describe the procedures that will be followed for the packaging and transport of samples from the field site to the laboratory for processing, including the labeling of such containers.

See Section *G.4.a.* Note samples will be shipped to other PACE consortium partners but these will be only pellets for qPCR that has been preserved in RNA later (which renders the cells non-viable) or AzCATI will have performed the DNA extraction steps on-site prior to shipment of the DNA. In addition, we may ship small amounts of biomass as freeze dried material, which also renders the cells non-viable. For this experiment, we will not ship off site any viable subject microorganism in any form.

G.4.c. Describe procedures for cleaning or disinfecting of planting, sampling, earthmoving and excavation equipment.

To ensure that the subject microorganism is completely removed from the test site after the experiment has been completed – all equipment (miniponds, trap ponds, and used sample containers) will be washed overnight in a 5% bleach solution, which has been shown to be sufficient to inactivate *C. sorokiniana.*

G.4.d. Describe procedures for the disposal of field test samples and materials that contain the subject microorganism strain(s).

Treatment of *C. sorokiniana with* 5% bleach or autoclaving is sufficient for inactivation. The inactivated samples will then be disposed of in the sewage system. Any bulk samples of biomass collected for future analysis (proximate composition or used in downstream conversion tests) will be freeze-dried, which also renders the cells non-viable.

Wastewater from the ASU Polytechnic campus goes to the city of Mesa's Greenfield Water Reclamation Plant (GWRP) located at from City of Mesa's website: the facility, originally constructed as a lift station, is a 16-million gallon per day water reclamation plant that treats sewage from southeast Mesa, the southeast portion of the Town of Gilbert, and all of the Town of Queen Creek. The facility is an activated sludge, nitrification/denitrification, "end of line" plant. It produces Class A+ treated water utilized for ground water recharge or ground water credits with the Gila River Indian Community (GRIC). The bio-solids from the process are Class B and used in non-edible land applications for agriculture.

G.4.e. Describe physical containment features such as the disposal of crop plants or dikes to contain water runoff.

The proposed physical containment of the field site miniponds is shown in Figure 10 and follows best practices established for outdoor secondary containment as practiced at AzCATI for closed cultivation system testing. Disposal of the subject microorganism is as described earlier. In the event of a catastrophic spill within the containment area, the entire area is flooded with 5% bleach solution and allowed to soak overnight. Surfaces (pond walls, stands, etc.) are also sprayed down with a hand sprayer with a 5% bleach solution then rinsed down into the containment area. Once the culture is neutralized, the area is sump pumped to drain and disposed.

G.4.f. Biological containment features of the field test (e.g. the use of trap plants) and the subject microorganism(s).

There will be no biological containment features included in the experimental design of the field trial because there is no known method to 'trap' algae species.

G.4.g. Describe access and security measures to be observed during the field test.

The AzCATI Testbed facility is completely surrounded by a fence that is **locked at all times** and access is strictly restricted to ASU staff and faculty pre-approved students. In addition we have CCD camera capability on the site and can monitor remotely when AzCATI personnel are not physically present (nights/weekends – though throughout the course of the experiment, the field site will have personnel on site for daily inspection/sample grabs even on weekends and holidays).

G.4.h. Describe containment procedures and training of personnel allowed on site.

All on-site workers will be trained to follow established SOP's. In addition, all researchers performing work with GM algae are required to take additional biosafety training course on the proper handling, storage, use in experiments, and disposal of transgenic microorganisms. Finally, the entire experimental protocol for this experiment will also be reviewed and approved by an internal, independent Biosafety committee at ASU as part of standard practice. All transfers for cultures/biomass between laboratory, greenhouses, and field site will be performed with proper secondary containment. The miniponds themselves will be within a secondary containment area as previously described.

G.4.i. Describe frequency and type of observations (e.g. ambient conditions or adverse effects) that will be made on site, submit sample observation forms.

See Table 9 in section G.3.c for weather effects to be recorded.

G.5 Describe waste materials handling and disposal procedures

G.5.a. Describe the application methods and precautions that will be followed to control dissemination during the initial release.

As previously described, all staff will be trained in the cultivation of algae and will follow established SOPs set forth by ASU EH&S. Strict labeling and recording procedures will be followed during the course of the experiment. The subject microorganism will always be appropriately labeled (Section *G.4.a*) and transported in leak-proof secondary containment. At the end of the experiment, all biomass will be disposed of as described in Section *G4.d*.

G.5.b. Describe any personal protective measures that will be followed to reduce human exposure.

Proper personal protective equipment (PPE) will be worn by all on-site staff (as required by ASU EH&S regulations). Proper PPE includes: gloves, lab glasses, lab coat, long pants, and closed-toe shoes. Each worker will also be limited in their exposure to the subject microorganism to no more than 8hrs/day during the lifetime of the experiment.

G.5.c. Describe waste materials handling and disposal procedures.

We will follow protocols that have already been established by ASU EH&S for the containment, labeling, and disposal of GMO algae. As mentioned previously in Section G4.d, waste from the experiment will be handled by bleaching or autoclaving the spent biomass.

G.5.d. Describe method, amount, frequency, and duration of application of microorganisms.

We will inoculate the miniponds at a total volume of 800-1000 L and a starting density of a minimum 0.1 g/L. The experiment will last up to 60 days. We intend to inoculate the subject microorganism only once during the experiment but if for some reason that initial application is not successful we will attempt to inoculate again after disposing of the biomass and cleaning the miniponds as described in section G.6.

G.5.e. Describe methods of cultivation after application if any.

G.5.f. Describe number of workers involved in application or subsequent activities, duration and routes of exposure.

G.5 e & F: The method of cultivation will consist of performing sampling on the miniponds, trap ponds and other AzCATI fieldsite ponds as outlined in Table 9. The primary monitoring and grab sampling of the ponds daily/weekly will involve up to 2 fulltime AzCATI staff with part time student worker support to monitor pH and temperature in the ponds, check on *in-situ* YSI sensors (for water quality monitoring), and acquiring the grab samples for further processing in the laboratory. Most samples will be processed on the same day, but some may take up to 5 days. For samples stored for OD/AFDW/N:P and scope, they will be stored and labeled as per Section *G.4.a.* and Table 9. Nutrients will need to be periodically added, in particular after pond harvest events and will be reset to target levels per the growth response of the ponds. Typical targets for nitrogen and phosphorus in the media will be 0.25-1X the values shown in Table 8. All other micronutrients are deployed at 1X what is shown in Table 8 at initial application and after any harvest/reset of the miniponds. For indoor sample processing as described in Table 9, again up to 2 full time staff supported by part time students will be involved in process the grab samples from the field. For harvest events, similarly, 1-2 staff supported by part time student workers. Supernatant from centrifugation for any proximate analysis sample or bulk harvest from the ponds will also be treated at 5% bleach level and discarded down the sanitary sewer per ASU EHS protocol. Routes of exposure and expected durations are as described in Table 7.

G.5.g. Describe worker safety procedures during application and cultivation.

Safety procedures will be followed as required by ASU EH&S.

G.6 Termination and Mitigation Procedures

G.6.a. Procedure to be followed after the field test is completed such as: level of the subject microorganism population at which no containment measures are necessary, use of plots after the field test is terminated.

Once the field experiment has been terminated, all biomass collected that has not been processed for analysis per Table 9 will be inactivated by bleaching or autoclaving. All equipment will be cleansed of the subject microorganism (including all sample containers, miniponds, etc.) by bleaching. Any pond spill will be contained within the secondary containment and treated with bleach, and liquid disposed of in the sewage system.

G.6.b. Define the type of unexpected effects and the quantitative level (if possible) which would necessitate the emergency termination of the field test.

Alongside of the experiment, we will utilize traps ponds that will be used to assess the dispersion capacity of the subject microorganism. These ponds will be placed at 5, 20, and 50 meters from the experimental miniponds along the path of the prevailing winds for that time of year. Additionally, we will monitor an additional set of traps at the 5 m range along the other major compass points from the miniponds. If we detect a 1000 fold increase of the subject microorganism by qPCR we will consult TSCA office for further action. As mentioned in Section G6.a, upon termination of the experiment, all biomass will be subject to bleaching or autoclaving. Additionally, all equipment will be bleached and/or autoclaved. If we experience any catastrophic failure (i.e. leak/draining of a pond) we will terminate the experiment, notify the TSCA office along with ASU EHS (ASU requirement) and perform root cause analysis. We will not restart the experiment without TSCA office and ASU EHS approval.

G.6.c. Describe the emergency termination procedures to be followed if adverse environmental effects are observed during the course of the field test.

If adverse environmental effects are observed we will consult the TSCA office to determine next steps. Upon termination of the experiment, all biomass and equipment will be subject to bleaching and/or autoclaving.

G.6.d. Describe how spills or leaks will be handled.

Any pond spill will be contained within the secondary containment treated with bleach and/or autoclaved and disposed of in the sewage systems.

G.7 Monitoring Endpoints and Procedures for Isolating Subject Microorganism

G.7.a. Relate the monitoring endpoints that will be evaluated to samples that are collected, such as population trends in soil or rhizosphere samples, or based on aerial dissemination during application as indicated by gravity plate samples.

As called out in Section G6.a, we will include trap ponds at various distances from the miniponds. Throughout the experiment we will monitor the trap ponds for the presence of the subject microorganism, which will be monitored via qPCR. We will also monitor other AzCATI field site open ponds via qPCR running independently of the TERA experimental protocol. The endpoint of the experiment is defined by the ecological experiments, which will be approximately 60 calendar days.

G.7.b. Describe comparisons between the subject microorganism strain and the unmodified parent that will be monitored in the field.

The comparisons between the recipient and subject microorganisms is outlined in the experimental objectives in Section G.3.a.

G.7.c. Describe the techniques used to isolate the subject microorganism from test samples and the rationale for this procedure. Include information on positive or negative controls used with sampling technique, if applicable.

The recipient and subject microorganism will be cultivated in separate miniponds. We will detect the presence of the subject microorganism in trap ponds and neighboring miniponds via qPCR analysis.

G.7.d. Describe the selectivity or specificity of the monitoring technique based on experimental observations under conditions similar to the field test site.

See Section *G.3.a.* for expected detection limits of PCR-based analysis in soil and media, respectively. Detection limits will be verified with the subject microorganisms prior to the start of the outdoor field trial.

G.8 Sampling Procedure

G.8.a. For each objective or monitoring endpoint identify the following: how, where and when samples will be taken.

Many different types of samples will be collected in service of the two objectives outlines in Section *G.3.a*. and refer to Table 9 (Section *G.3.c)* for where and when samples will be taken.

G.8.b. Describe how the samples will be labeled so that they can later be traced back to their source.

Each sample will be labeled as described in Section *G.4.b.*

G.8.c. Include the standard procedures for preserving, processing, and analyzing samples.

All SOPs are attached to this TERA submission in a single pdf file (SOPs.pdf)

G.8.d. Describe methods of measurement, equipment used, the precision and accuracy of the method of analysis

G.8.e. Describe methods for the statistical analysis of field data.

G.8 d & e: The primary metric we will be looking to evaluate for this experiment is difference in biomass productivity/growth rate. Samples will be collected from biological replicates (n = 3) for each of the subject and recipient microorganisms and samples processed as technical replicates in triplicate AFDW, OD and N:P in the media. Biomass productivity and nutrient uptake (consumption) rates will be compared between subject to recipient in order to resolve a 20% or greater difference between the strains via t-test (alpha value = 0.05). Other metrics will be evaluated (e.g., differences in proximate composition) and will employ the same statistical methodology. A power analysis will be performed to determine minimum number of samples for P = 0.80 level for a given metric being evaluated.

G.9 Record Keeping and Reporting Test Results

G.9.a. Describe the frequency of reports on the field experiment and the proposed format of the reports.

We will produce a final report at the conclusion of the experiment. The report will contain the following sections: Executive Summary, Primary Experimental Objectives, Background, Experimental Design and Analysis, Deviations and Observations, Results/Summary, Conclusions/Discussion, and Appendices/References. The final report will require approximately 3-4 months from the end of the field trail to complete, although we expect to produce interim/preliminary results and those will be made available as practical.

G.9.b. Outline the contents of the progress reports including a summary statement, statistical analysis procedures and presentation of the raw data.

Any interim progress reports will include a brief summaries of the results to date, any deviations in protocol. The main vehicle for presenting preliminary results will be in quarterly reports to the project sponsor (DOE) and in technical presentations to the PACE consortium.

G.9.c. Describe procedures for filing raw data and information on procedures followed for the analysis of the samples.

Data will be recorded in lab notebooks, batch traveler logs, instrument logs, as well as compiled in a standardized spreadsheet log each time data is collected and samples processed. Data in notebooks will be signed and countersigned the same day that data is collected. Raw data will be filed and stored on the AzCATI's shared drive as well as archived on the PACE Consortium shared drive in Dropbox.

G.9.d. Describe compliance with Good Laboratory Practices.

This R&D experiment will not be done according to GLP requirements. However, data will be collected with dual signatures, data and samples will be archived, reports will be archived, weather conditions will be monitored and recorded, names of scientific personnel will be recorded, and any deviations from the experimental plan will be recorded and the impact to the study described.

References

ANL; NREL; PNNL. (June 2012). Renewable Diesel from Algal Lipids: An Integrated Baseline for Cost, Emissions, and Resource Potential from a Harmonized Model. ANL/ESD/12-4; NREL/TP-5100-55431; PNNL-21437. Argonne, IL: Argonne National Laboratory; Golden, CO: National Renewable Energy Laboratory; Richland, WA: Pacific Northwest National Laboratory.

Bagchi, D., Preuss, H.G., Swaroop, A., (2016), Nutraceuticals and Functional Foods in Human Health and Disease Prevention. CRC Press, Taylor & Francis Group. Boca Ratan, FL. (Edited) Chapter 39: 673-682.

Barry, A., Starkenburg, S.R., Sayre, R. (2015). "Strategies for optimizing algal biology for enhanced biofuel production and carbon capture." Frontiers in Energy Research 3(1): 1-5.

Becker, E. (2007). "Micro-algae as a source of protein." Biotechnology Advances **25**(2): 207-210.

Bioenergy Technologies Office Multi-Year Program Plan (2016) Washington D.C, 2016.

Bock, C., Krienitz, L., Proschold, T. (2011). "Taxonomic reassessment of the genus Chlorella (Trebouxiophyceae) using molecular signatures (barcodes), including description of seven new species." *Fottea* 11 (2): 293–312.

Gonzalez-Esquer CR (2018) "Nuclear, chloroplast, and mitochondrial genome sequences of the prospective microalgal biofuel strain Picochlorum soloecismus" Genome Announc 6:e01498-17. https://doi.org/10.1128/genomeA.01498-17.

Henley, W.J., Litaker, R.W. Novoveská, L., Duke, C.S., Quemada, H.D., and R. T. Sayre, (2013) "Initial risk assessment of genetically modified (GM) microalgae for commodity-scale biofuel cultivation" Algal Research 2: 66-77.

Lemieux, C., Otis, C., Turmel, M. (2014). "Chloroplast phylogenomic analysis resolves deep relationships within the green algal class Trebouxiophyceae." BMC Evol. Biol.14 (1): 211.

Laurens, L. et al (2017) IEA Bioenergy, State of Technology Review - Algae Bioenergy (ISBN: 978-1-910154-30-4)

Liu, J., Gerken, H., and Y. Li. (2013), "Single-tube colony PCR for DNA amplification and transformant screening of oleaginous microalgae" J Appl Phycol 26:1719–1726

Metting, B. and J. W. Pyne (1986). "Biologically active compounds from microalgae." Enzyme and Microbial Technology **8**(7): 386-394.

Neofotis P., Huang A., Sury K., Chang W., Joseph F., Gabr A., Twary S., Qiu W., Holguin O., and Polle, J.E.W. (2016). "Characterization and Classification of Highly Productive Microalgae Strains Discovered for Biofuel and Bioproduct Generation." Algal Research 15: 164-178.

Quinn, J., Davis, R. (2015) "The potentials and challenges of algae based biofuels: A review of the techno-economic, life cycle, and resource assessment modeling, Bioresour. Technol. 184 444–452.

Rajamani, S., Teplitski, M., Kumar, A., Krediet, C.J., Sayre, R.T. (2011). “N-Acyl Homoserine lactone Lactonase, AiiA, Inactivation of Quorum-Sensing Agonists Produced by *Chlamydomonas reinhardtii* (Chlorophyta) and Characterization of *aii*A Transgenic Algae.” Journal of Phycology 47: 1219-1227.

Ramírez-Romero, R., Rodríguez-Tovar, L.E., Nevárez-Garza, A.M., López, A. (2010). “Chlorella Infection in a Sheep in Mexico and Minireview of Published Reports from Humans and Domestic Animals.” Mycopathologia 169:461–466.

Ratha, S.K., Prasanna, R., Gupta, V., Dhar, D.W., Saxena, A.K. (2012). “Bioprospecting and indexing the microalgal diversity of different ecological habitats of India.” World J Microbiol Biotechnol 28: 1657–1667.

Roberts, A., Haighton L.A. (2016), “A hard look at FDA’s review of GRAS notices.” Regulatory Toxicology and Pharmacology 79: S124-S128.

Siripornadulsil, S., Traina, S., Verma, D.P.S., Sayre, R.T. (2002) “Molecular Mechanisms of Proline-Mediated Tolerance to Toxic Heavy Metals in Transgenic Microalgae.” The Plant Cell 14: 2837-2847.

Soo, R., Woodcroft B., Parks, D., Tyson, G., Hugenholtz, P. (2015) “Back from the dead; the curious tale of the predatory cyanobacterium Vampirovibrio chlorellavorus” PeerJ 3: 968-990.

Tischnera, R., Planchet, E., Kaiser, W. (2004) “Mitochondrial electron transport as a source for nitric oxide in the unicellular green alga *Chlorella sorokiniana”* FEBS Letters.